FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTCAGATCTGCTCGGTAGA CCTGGTGCACCACCACCATGTTGGCTGCAAGGCTGGTGTCTCCGGACACTACCTTCTAGG GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA GÁACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAAATATTTAAAATT GATCAGATGGGAAGATGGTTTGTTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCCTCTTCTCATCAGAGCTGCATGGTACAC AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT AGTTCTTTTCAGCATGTTCCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAATGGGGCCAGATATGC ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT TTGGTGAATGTGAAAACTAAAGTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTTGCATATTTTTTTGGAGT GCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA GCTGAACTTAACAAAACTGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG CTTCTCAGTGCTCTTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTTACTTTTGAATGTTACAAAAGGAA ATAACTTTAAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG AATACAAACAGTATACTCATG

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYYGLGLSNEIGAIEKAVIWPQYVKDRI HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFSMFLLYDTQKVIKRAEVSPMYGV QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

FIGURE 3

GCCGTGCCCTTGGTCGGCAAACTCGGCCTCATCAGCCCGGCCTACCTCTTCCTCTGGCCCGA AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG GTCCAGGAACTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA GATACCCAATGGACTTGGGAGGAAGAAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG CTGCCAGTAGGAGAGGAGTATCAGGATTTGGTGTGCCCCTGCTAGCATGAGGCGAGC TGCTGATCAGAATGGCGGAGGCGGGAGACACAACTGGGGCCAGGGCTTTCGACTTGGAGACC AG**TGA**AGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCCTCCCAGTGCTGGGTG CACTTAACAACTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCCACAAGTTTCACGAT TCTCATTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAATTGCAAAACTGAC ACAACAATCATATTCACGTTATTTTCCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCCC TTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTTGCTTTGTGT TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACTTCTCATTTTGAATAGAGTTTGTGTG TTTTTGTAAACTAATCCTTTTTATTGGTAAAAATTGTAAATTAAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

FIGURE 5

GGGGCCGCGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT CCTTGTGGCCCAAAGGCCTAACCGGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC CCTTTGGGGCGGGATGCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG CGGGGTTCCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA AAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTACCT GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAAGAAAAA AACAGTTATCAGAGGCTAAAACAGAAGAGCCCACAGTGCATTCCAGTGAAGCTGCAATAATG CACAAAAAGGCCTGAAGATTCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAG**TA** CTTACACTG

FIGURE 6

MAAEEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVVSDLEHEEMKILREVL RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEEKQTLLKRRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATTGACCAGGAA

FIGURE 8

TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGGAGGTAATTAAAAAAAC AGTGGA**ATG**GAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT CTGCCAATGAAGAAACAAGTATGATTATCTTCCAACTACTGTGAATGTGTGCTCAGAACTG AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTCATGAAGTGGTCCATTCCTGCCT TTCTTTATTTCCTGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATG GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAGGATAGTGCTGAA GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCT TGACTGCCGGGACTAAAACTTTACAGCACAACTTGGCAGGACGTGGATTTCATCACGATGCC TACAGCAAAGGAATGGACTTTTCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTCAGTC ACATCCGTCTTGGCATGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA GAACAGCAAACTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTCAGTAGCC CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA TTTATTTATAATGCCAGCAAGCCTCAAGTTCCGGAATACGCACCTAGGCAAGAAAGGATCCG AGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTC**TAA**CTGGTACCCACATAGTTTGCA GCTCTCTTGAACCTTATTTTCACATTTTCAGTGTTTTGTAATATTTTATCTTTTCACTTTGATA AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT TTTCCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT CTTCCCTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT TTCTCAGACACACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC CAGATTGTCAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAAGGATCACTT AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTCATG GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT CATTCTCAGAAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA AGGTAATATACTATTATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG GTGGAAATTTGTAATTAAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV IFSNFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI YNASKPQVPEYAPRQERIRDLSGNLWERSSGDGEELERLTKPKSDESDEDTF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293, 298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

FIGURE 11

CGGACGCGTGGCCGTGGGCGGACGCGTGGGGCCGGCTTGGCTAGCGCGCGGCGCCC GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG AGCGGCTCCGCGGGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA GGTGAAGAAGAACCGAAAAAGAAGAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG • CTATTGGATGTGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC CTGGGATGCCATCACAGACCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC TGGGTCGCCTTATGCCCTGGATCATCTTCTCCACGCCCCTGGCCGTCATTGCCTACTTCCTC ATCTGGTTCGTGCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTCATCAGCA ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCACTTCACAC AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGTCATTGTCTGTATCTATATAATCTG TGCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGACCCTATGAAGCCCAGCAGTCTG AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTT ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC TACCCTGGTCCATGCTGCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTCACCAAGTTTGCCTCTGGAGTGTC ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCAGAAACAGACTCCACAG AGCTGGCTAGCATCCTC**TAG**GGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG TGGCCTCCTGCCTCCCTCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG SCPTSHTARPIGTCFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV ILILGVREQREPYEAQQSEPIAYFRGLRLVMSHGPYIKLITGFLFTSLAFMLVEGNFVLFCT YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKKAL QALRDEASSSGCSETDSTELASIL

FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATG**T GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT TTCATATTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG CAGTTTTATGCATTGCTACCATTTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA GAGAACGTTATCATCAAATTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGG ACTTTCTATTGTGGCAAACTTCCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG CAGCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG AGTAAGTGCACTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG ATTTAGAACAGAAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATT**TGA**TGAAAGGAT AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTCACAGAAGTTGCTTA TTCTTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI AAVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG AVLTFGMGSLYMFVQTILSYQMQPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYD TAPCPINNERTRLLSRDI

FIGURE 16

FIGURE 17

CCCACGCGTCCGCCGCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGG CCGGGGTGCGGAGCCGACATGCGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC AGGCTGGAGGCAGGTCGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCCTGCTCTTCTGCGGCGCCCTA CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT TGCTACCTGCTCCCAGTATTTTTGGCAAACAGTTGGTGGTGTCCTACTTTCCTGATAAAGT TGAGACTTTTCCCCATGACACCAAACTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT CCCATCGTGCAGTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTCATCTGTGT GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTGGATGCTCTTTTCTCCTGGGACACTG TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAAATTT AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA CACA**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT ACAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYR KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS IFGKQLVVSYFPDKVALLQRKVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPILNIPIVQFF FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ LNETSTANHIHSRKDT

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGGAGGGCCCGAGGGGGCCCCGCATGAATCATTGTAGTCAATCATTTT CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA TAGGAAAATAACTTGGGATTTTATATTGGAAGAC**ATG**GATCTTGCTGCCAACGAGATCAGCA TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC CCCCCGCAGTATCCTCTCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA CAGACTTTGACCCCTGGTGGÄCAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCCTGCC AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCCTTATCCATGGAGGAGACCTCTGAA CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGGAGTAAGATGCATAAG ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTTCCA GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCG ACACCACCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT $\tt GGAACCGCTTTCTCAGAACTG{\color{red}{\textbf{TAG}}}GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA$ AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

FIGURE 20

MDLAANEISIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY KVLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPFPYPWRRPLNRSQMLRELFPV FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRHCQSVAMP IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

FIGURE 21

FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTGGTGCACTGGTGCCACTGGTGCCACGCTGCTAGACCGTGCCTATGAGCCGCTGCGGGGCTGCAGTGGGGACTGCC CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAACC**ATG**TCCCAGTTCGAAATG GACACGTATGCTAAGAGCCACGACCTTATGTCAGGTTTCTGGAATGCCTGCTATGACATGCT TATGAGCAGTGGGCAGCGCCCAGTGGGAGCGCCCCAGAGTCGTCGGGCCTTCCAGGAGC GTCCACCTGCGGCGTTTCAACCTGCGCCGTTCAGCACTTGAGCTCTTCTTTATCGATCAGGC CAACTACTTCCTCAACTTCCCATGCAAGGTGGGCACGACCCCAGTCTCATCTCCTAGCCAGA CTCCGAGACCCCAGCCTGGCCCCATCCCACCCCATACCCAGGTACGGAACCAGGTGTACTCG TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT GCTGCGTGCCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACTTCGAGTACT TGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTC CCTGGGTCTGCAGGACTACGTGTCCCAACCTGGACCTGTCTCAGTACCCTGTGTCCGCTGGTCTTCAGGACTACGTGTCCCAACCCTGGACCTCAGCAACCCAGCCGTCTTCCGGGACCTGTCTAAGCCCATCGGTGTGGAACCCCAAGCATGCCCAGCTCGTGAGGAAAGCTTTGAGGACCCAGCAGGGACCATTGACAAGTTCCACTATGGCACCACTACTCCAATGCAGCAGGCGTGATGCACTACCTCACCGTGGAGCCCTTCACCTCCTGCACGTCCAGCTGCACGTCCAAAGTTGCACAAGTTGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGCAGCCTGGC AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCCT GACTTCCTGGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT AGGCGATGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACCGCC AGGCTCTGGAGTCGGAGTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGC TACAAGCAGCGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA TCAGCAACTTTGGGCAGACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACTCGGCTCTCA GCTGAGGAAGCAGCCCATCGCCTTGCACGCCTGGACACTCACCTAGCATCTTCCAGCA CCTGGACGAACTCAAGGCATTCTTCGCAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT ACCATGGGCAGCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGTGAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAAGCTGCTATTCAGCGGTGGCCACTGGGATGGCAGCCTGTGCAGCTGCCACCCTTGATGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCGGGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTCAGTAGGCCTGGCACCAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTCCTGCAGTGAGCTGTGTGGCCATCAGCACT GAACTTGACATGGCTGTGTCTGGATCTGAGGATGGAACTGTGATCATACACACTGTACGCCG GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCTGGAGGA TGGCAAGCTCATCGTGGTCGCGGGGCAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGCGGAAGCTGTGGCGGTCCTCGCGGCGCATCTCCCAGGTGTCCTCGGGAGAGACGGAATACAACCTACTGAGGCGCGCTGCTGGCCAGTCCGGCCCCCCGGCAGGCCTG ${\tt GCCCGGGAGGCCCCG}$ ${\tt GCCCGGGGGAACACCCCGGGGTGGCCAGCCCAGGGGGTGA}$ GCGGGGCCCACCTGCCCAGCTCAGGGATTGGCGGGCGATGTTACCCCCTCAGGGATTGGCG GGCGGAAGTCCCGCCCTCGCCGGCTGAGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL RYTAVLKOOATOHSMALLHWGALWROLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAELETP MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP LAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELIP EFFYFPDFLENONGFDLGCLOLTNEKVGDVVLPPWASSPEDFIOOHROALESEYVSAHLHEW IDLIFGYKORGPAAEEALNVFYYCTYEGAVDLDHVTDERERKALEGIISNFGOTPCOLLKEP HPTRLSAEEAAHRLARLDTNSPSIFOHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF SFSKDPTMGSHKTORLLSGPWVPGSGVSGOALAVAPDGKLLFSGGHWDGSLRVTALPRGKLL SOLSCHLDVVTCLALDTCGIYLISGSRDTTCMVWRLLHOGGLSVGLAPKPVQVLYGHGAAVS CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGOPSEVRSSOFARKLWRSSRRISQVSS GETEYNPTEAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697, 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840, 873-879, 912-918, 954-960

FIGURE 24

CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC GCACACTCCGTTACCACACTGGGTCATTGGCATTTGGAGCCCTCATCCTGACCCTTGTGCAG ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTGGTGGGGGGTCCTGTCCTTCTTT TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC GAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC ACCCCACCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG AGGCGGCCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC TCGGGAGGCTGAGGCAGGAGATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA AAAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKN
EAPPDNKKRKK

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT GGCTATGTTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT CAGGAAATGACAGTGATGGGTCAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA GTGGAGCAAACCATGCGGAGGAGGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCCTGCAGCG CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA GCGGCTCCAGGAGTTCCTTGCAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGGAAATGATTGAAGAGTCTGCAAATAAA TCTGGCCAGCGACGTGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC CTTTGCACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA AGTCCTTTGTGTGTTCGACAAGAACCGGCGCTGCAAACTGCTGCCCCTGGTGATGGCTGCC CCCCTGAGCATGGACATGGCACAGTGACCGTGGTGGGCATCCCCCCAGAGACCGACAGCTC GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT $\mathtt{CTGGACGCACTTATTTCCCTCCTGTCCT}$ TATGTAACTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT TTTTATTAAATAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAYEDIFRDEEEDEEHSGNDSDGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEYHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSLCNTSYTAARFKLWSVHGQKR
LQEFLADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTFSIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLSRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFL
DALISLLS

FIGURE 28

FIGURE 29

CAGGAACCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC GAGTAAAGTACGCTCCGGTCACC**ATG**GTGACAGCCGCCCTGGGTCCCGTCTGGGCAGCGCTCCCTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC CAGCGGCTGCCAACGGTGCTGTGACTCTGAGGACCCCCTGGATCCTGCCCATGTATCCTCAG CCTCTTCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC CATGGCGACCGGCCAGTTTGCTGCTCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCATC
CTGTACGCGCAGCCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGC
CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA TTTGCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTCTGGGGAGA GCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCCAATTCCAACATTCTTAAATTCTCTCAT GTCCAGCCCAGGCGGGAGAGATGTGTACATAGGTTTTAAAGCAGACCCAGAGCTCATGGGGGCTGTGTTCTGGGTGTTCAGGTGCTGCTGGTCCTCCATTACCCACTGCTCCCAAGGCTGG TGGGACGGGTCCCGGTGGCAGGGCAGGTATCTCCTTCCCGTTCCTCATCCACCTGCCCAGGCCAGGCCAGGCCAGGTCATCGTTACAGCAAACCCCAGGGGGCCTTGGCCAGGTCAAGGGTTCTGTGAGGAGAGAACCCAGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA TAGCTCTCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCCCTTCCTGCCTCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTCCACCAAAGTCATCT CCATTGGTGCTCATGCAGACTCTGGGGGCTGAGGTGCCCCGGGGGGGTGATCTCTGGTGCTCAC

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSGRPH ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET YVHIMHNQKEAVILYAQPSERSIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCGACCCGCCAGGAAAGACTG AGGCCGCGCCTGCCCCGCCCGCCTCCCTGCGCCGCCGCCTCCCGGGACAGAAG**ATG**TG CTCCAGGGTCCCTCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT CACCATGCTCGACGCAGCCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC AGAACCAGATCGCCAGCCTGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC CTGCCCGCCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCCCTGCGGCTGGCCGGCAACAC CCGCATTGCCCAGCTGCGCCCGAGGACCTGGCCGGCCTGCCCTGCAGGAGCTGGATG TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG CTGCTGCCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCTG GGTGCGCGAGAGCCACGTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAGCCACCACC ACCACAGCCACAGTGCCCACCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCTTAG CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCCACCGTCCACCTGCCTC AATGGGGGCACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTT CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGG CCCGGGCGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGCCCATACACCCCCAGCCGTCCA CTCCAACCACGCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCG CCCTGGCCGCGTGCTCCTGGCCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG GGGCGGCCATGGCAGCGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCT GGAACTGGAGGGAGTGAAGGTCCCCTTGGAGCCAGGCCCGAAGGCAACAGAGGCGGTGGAG AGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCCTTCCCAGGGCCTGGCCTC GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCCAGAAC CGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCCTGCGGGCTCTCCCACTCCAGGCGGA CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGGAAAGGAAGATGCTTTA GGAACATGTTTTGCTTTTTTAAAATATATATATTTTATAAGAGATCCTTTCCCATTTATTCTG GGAAGATGTTTTCAAACTCAGAGACAAGGACTTTGGTTTTTGTAAGACAAACGATGATATG AAGGCCTTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVPLLLPLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFEN GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP GILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAG NTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG PWVRESHVTLASPEETRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALS SSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLRL TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPA VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPGAG PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**AGACTCATCAG TTTTCATTCTGTCTCCAAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG GTGGGGCAAAAATACAAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAACAAACTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTTGCGTGATGGAATCA TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC TCAAATACTGTAATGAGAACTATAAAATTGGAGCATGTACATTTCAGAGTGTTTTACATTCA ACAGGATAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATTCCAATATTTAAATTTTGCC AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACTGCCTCACTTGAAAACTCT CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC CCTTGGAACACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAAATTGCTCA ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAACTAG TTACTGGAACCCATTCCATTCTATTGCATTCCCACCAGGTATCATAAACTGAAAGCTCTCCT GGAÁAÁAAAÁGCATACTTGGAATGGCCCAAGGATAGGCGTAAATGTGGGCCTTTTCTGGGCAA ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT A**TAA**AATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA CAACCTTTATGATGGCAATTTGACAATATTTATTAAAATAAAAAATGGTTATTCCCTTCATA TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG GCTTATGGAAAAAGGTGTTCATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCCTGGTATGGAAAAAACATATTAATAT GTTATAAACTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTGCCATTGAAATGGCATTGAA ATAAAAGTGTAAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTTGGA TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTTAGAATGA ATCTGTATTTCTTTTATAAGTAGAAAAAAAAAAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQLDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNDENCSWPETVVNMNLSYNKLS
DSVFRCLPKSIQILDLNNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVLGLAVAFCCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHHNLFHENSDHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTCGCCGGGAAAGG GAGGGAAGAAGGAAGGCCGGCCCGCCCCTGCGCCCCGCGCCCTCTGCGCCCCTGTCCGCCCCGGC CTGCTGTGCCCTGCGCCCTTGCCCCGCGCCAGCTTCTGCGCCCGCAGCCCGGCGCCCCCGGTGACCGTGA $\tt CTGGCAGTGACCCTGGCCGGGGTCGGAGCCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT$ CCGGGGAGGGTGGGAGCGCCCCCGCAGGAGCCCAGGCCCCAAGAGGCCCAAGCCCAAGAAAGCTCCC AAGAGGGAGAAGTCGGCTCCGGAGCCGCCTCCACCAGGTAAACACAGCAACAAAAAAGTTATGAGAACCAAGAG TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA CATCGAGGGAGACTCAACATCCAGGCGGGCATTAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAG AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCATCACTCAAGGGA GGAACTCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTC ${\tt ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT}$ ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGA GAATGGAGATCCTGGGCTGCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT CAATATCACCAGAATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCCCACGGCAATGAGGTGCTGGGC CGGGAGCTGCTGCTGCTGCTGCAGTTCGTGTGTCAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT GGTGGAGGAGACGCGGATTCACGTCCTCCCTCCCTCAACCCCGATGGCTACGAAAGGCCTACGAAGGGGGCT CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACTTTCCTGATTTA AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCCAATCACTATATTGCAATCCC CTTTTGTGCTGGGCGGCAACCTGCAGGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCGGTCCCCC ACACCGCCTCATGACAGACGCCCGGAGGAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGGCACTGTCA ATGGGGCCTCCTGGCACCCGTCGCTGGAAGTCTGAACGATTTCAGCTACCTTCATACAAACTGCTTCGAACTG TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATC TCTGATCGTGTTCATGGAGCAGGTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTCACATGGAAAAGGAATCC CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTC CTGAACCCTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG CTATGACATGGGGGCCACAAGGTGTGACTTCACACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGG AGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGCCAGCCGAGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGT TGGACTCACTCACTGTTGTTTCCTCTGTAATTCAAGAAGTGCCTGGAAGAGGGGTGCATTGTGAGGCAGGTCC CAAAAGGGAAGGCTGAGGCTGTTTTCTTTTCTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCA GCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAGCCAACCTGGGAATCAGAGAGAAGGAGAAGGAGAAGGAGGAGA GAGCCTGTCCGTTCAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGCGTGGCAGCAAGG GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAAATGTTACCA TTTGAGATGCTCCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCT TGGAGCTTCTTGCACAAATTCTGGGTCCATAAACAACCCCCAAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT $\tt CCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC$ GAATTGAGTGCTCATGGGTTGGCCTCATATCAGCCTGGGAGTTATTTTTGATATGTAGAATGCCAGATCTTCCA GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGAALEDPDYYGQEIWSREPYYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSYAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSKTNMARIREIMEKFGKQPVSLPARR
LKLRGRKRRQRG

FIGURE 37

 $\tt CTAAGAGGACAAG{\color{red} {\bf ATG}} AGGCCCGGCCTCTCATTTCTCCTAGCCCTTCTGTTCTTCCTTGGCCAAGCTGCAGGGG$ ATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC TTCAGCTCCAGCTCCAGGTCGGGCTCCAGCTCCAGCCGCAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT TTCCAATTTCACCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAATTGACATCAT GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAAC TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTCGCCGAGAAATCGTGGC TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCTCCTCCCACTC CAGGGAGCTGTGGTCATGGTGGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAGGACTGTATTGGGTGGCGCC ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACTCTCCCTAA TGCTGCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT ${\tt CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTTCT}$ GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGG GCAAACTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACTATAACCCTTTTGACCAGAAA $\tt CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAG{\color{red}{TAA}} GCTGTTTA$ $\verb|CTAAAAGTGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT| \\$ TTGTCAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA GGAATTAAGGAACTTAAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC GGAGCTCCTCGAGGGACCAAATCTCCAACTTTTTTTTCCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACTTGTAAGAAAAT GAAAATTTAATTTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT

FIGURE 38

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGTGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

FIGURE 40

CCGCTGCTCTTGTGACGTTGTGGAGATGGGGGGGGGGGCTCCTGGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAACTCCACTGTAACTAGATTGATCTATGCACTTTTCTTGCTTTGTTGGAGTATGTAGCTTG TGTAATGTTGATACCAGGAATGGAAGAACAACTGAATAAGATTCCTGGATTTTGTGAGAATG AGAAAGGTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTTGTGCTTT GAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGAACTCGAGATGTTGGTATGCAGCCTTGT TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC TACACTCATCCAGCCAGTTGTTCAGAAAACAAGGCGTTCATCAGTGTCAACATGCTCCTCTG CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC TGTCCCAAAGGAAGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC TCTTTTTGTTGTGTGTATTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA CTGACTCTAACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC AATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCACTTGTTC TAGTAAAAGTGGCCATTATTGGGCTTATTCTCTGCTCTATAGTTGTGAAATGAAGAGTAAAA GCAAATGTATGGCTGCCTTTTGAAATATTTGATGTGTTGCCTGGCAGGATACTGCAAAGAAC ATGGTTTATTTTAAAATTTATAAACAAGTCACTTAAATGCCAGTTGTCTGAAAAAATCTTATA AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG TTATGGATGGAGGTGTCGGTACTAAATTGAATAACGAGTAAATAATCTTACTTGGGTAGAGA TGGCCTTTGCCAACAAAGTGAACTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTCAGT GGAAATGTTTGCAACAAAGTGAACTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTCAGT GGAAATGTTTTGGAACTCTGAAGGATTTAGACAAGGTTTTTGAAAAGGATAATCATGGGTTAGA AGGAAGTGTTTTGAAAGTCACTTTGAAAGTTAGTTTTTGGGCCCAGCACGGTAGCTCACCCTT GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA GCTTGGCACATGGTGAACCTGTTCTATAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCCAAAAAGTTATAAAAATCA AGGCAAAATTTTGACAGGGAAGGAAGTAACTGCAAAACCACTAGGCTTTAGTAGGTACTTAT ATAAAATCTAGTCCAGTTCTCTCATTTAAAAAAATGAAGACACTGAAATACAGACTTAAATA GCTCAGATAGCTAATTAGGAAATTCAAGTTGGCCAATAATAGCATTCTCTCTGACATTTAA AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTAAAATGCTTTTGTAT TCTTCTTTAAAAATATTTATTAATGTGAATGGAATATAACAATTCAGCTTAATTCCCCAACC TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTTACCTCTAAATAA ATGAATTCAGAGAAAAAAAAAAAAAAA

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVWVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAANTCCACTGTAACTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGAAT
GTTGATACCAGGAATGGAAGAACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

FIGURE 44

FIGURE 45

FIGURE 46

GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGGCAGTCGAGTGTTTGCAGACCGGGGCGAG GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC CCCTGGTTTGTGTTTACGCACACACGTGCACACAGGCTCTGGCTCGCTTCCCTCCTTCTTTCCAGCTCC TGGGCGAATCCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTG AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGAT TGGAAGCTCGGCCTTCCTGTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCA ACATCATCCTGGTGCTGACGGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC ATCATGGAGCAGGGCGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCTCACGCTCCTC CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC AGGCACAGCACGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTCGGACTCCTTAAAAACTCCCG CTTTTATAACTACACGCTGTGTCGGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA CAGACCTCATCACCAATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCCTC ATGGTCATCAGCCATGCAGCCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCTCTTCCCAAACGC ATCTCAGCACATCACGCCGAGCTACAACTACGCGCCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCCTCATGTCGGTGGAC GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG TCCCGTTCTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCCTCAACATTGACCTG GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG AGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTAC CAGCGTGTGAAGGACCTGTGTCAGCGTGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC AGCCTGGCCGGACGCCGGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG $\verb|CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCCAGCCCCGAAACCTCA||$ CCAAGCGGCACTGGCCAGGGGCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC CTTCCCGACTACTCAGCCGCCAACCCCATTAAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA CCCTGCAGAACAAATTAAGAACCTGAGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCCAGAAGAATGTGAC TGTCACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAG GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAAAACTCCGCAAGCTGC TCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT GAGGACCATCAATGAGACTCACAATTTCCTCTTCTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCA ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACACGTACAG AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG ${\tt GACAACTGTGGGAAGGCTGGGAAGGT}{\color{blue}{\textbf{TAA}}}{\tt GAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA}$ CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC ATTTTTGCCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC ${\tt TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTTATTCTT$ TTGGTTTGTCACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCCTGGTTGCCTCTGAAGAAA CTGCCTTCATTGTATATGTGACTATTTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTAATAAGAAAT CCCAATTTTCAGGAGTGGTGGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHES
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVOLMELRSCKGYKOCNPRTRNMDLDGGSYEQYROFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGGAAAAGGAGTGAGGA GCTGCTGGGCAGAGGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCCACC TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC CGCTCGCTGAGGCTGCTGTCGCCGGTGCCTGTGGACAGCAGCTGCCCCTGCCCTCCCATCTG TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT

FIGURE 49

MLGLLGSTALVGWITGAAVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRAQPWPFR RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPHRHHPRHAR

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCC**ATG**TCGGACCTGCTA CTACTGGGCCTGATTGGGGGCCTGACTCTTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG GTACTCAGGGCTACTGGCTGGGGTGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCA CTGTGGCCTACAAGTTCCACATGGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG AGATCTACCAGGAAGACCAGATCCATTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT GTGCCTGAGATGAAGGAGACAGAGTGGAAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGGAGTCACGGCTGGACCCTGGGACTGAGC CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCCAAGGAG**TAA**CCC ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT CCAGCCCTCTTCCTCCTCTGGGGGAGGAGGGGTTCCTGAGGGACCTGACTTCCCCTGC TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA GGGACTATTTCTGCACCAGCCCCCAGGGCTGCCGCCCCTGTTGTGTCTTTTTTTCAGACTC ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAA AAAAAAAAA

FIGURE 51

MSDLLLGLIGGLTLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPLAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGPLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

FIGURE 52

FIGURE 53

MTLRPSLLPLHLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS KKKLKEEKRNKSKKK

FIGURE 54

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAAAACGAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG TGAGGCGGGCGGCGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG ACCTGAAAAAA**ATG**TCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT GTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG CTAAAGAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAG**TGA**ACACATCTGAT TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACTT ATAAAATGTTAGAGGAAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT TGTAAAAATAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAAATATTCCGTGG TCAAAATTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT GTAACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC CACATCCACCACTG

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWWIIIDAAVIYPTMKDFNHSYHACGVI ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

FIGURE 58

FIGURE 59

TGGACGGACCTGAAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTCCAGAATGCC

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

FIGURE 62

FIGURE 63

TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC AAACGGCCCCAGCGCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC TTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGCTGGAACAACAT ACCAAGGCAGGGGGGCTGCTGTCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGAGAA CCATTACATTGGGCGTATGGAAGAGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG AAGGGCAGCTGGACTGGGCTCCCCTGTCCTCTCCTTTTGACATCATGGTACTGGAAGGGCCC AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTCAGGGCCTCAAGGA GAAGTTTCCACAGGAGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA GTGGAGCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCGTGGTTCAGCTCCTCGAC AGGTGTGGGCTGCTGACTCCTTCCATTCCTTCAAGCATCCACCCAGAGCCTGGCTGA AGGTGTGGGCTGACTCGTTTCTCTCCATTCCTTCAAGCATCCACCCAGAGCCTGGCTGA
GGTCCTGCAGCAGCTGGGGGGCCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCA
CTTACGGTGTCACCCCCAACCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC
ATGAAAGGAGGCTTTTATCCCCGAGGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT
GATTCAGCGGGCTGGGGGGCGCTGTCCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT
CAGCTGGGAAAGCCTGTGGTGTCAGTGTAAAAACCCTATGAACACCTACTGCCGGGAACGC
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGAACGC CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGCGGCCCGGCTTAGGCATGACCT CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAG**TTCCATCAGGGAGG AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG CATTAGTTCCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG TTTAAATCACAATTCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTCAGTGGCTCTTCAGGGGACAGGAAAT GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTTGCAGTTAATTAGCACATGTGAGGG TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTGGTATCAGACATACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTG ATTCCATTGCTTTAAAAAAAAAAAAAAAAAAA

FIGURE 64

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ EEAIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMPREEAAEH IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEEWQAELKGKRGSDYETFKNSFVEA SMSVVLKLFPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

FIGURE 65

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA GGGGTTGGCACCGGCCCCGAGAGGAGGATGCGGGTCCGGATAGGGCTGACGCTGCTGTG TGCGGTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT GTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGAATTAGAATCCTCTATTCAAGA AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTCACAGAAGATATCAGCTTTCTAG ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCCTAGATAA GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTTGTGAAACTGAAGAAGAGGCTGCTAAGAGA CGGCAGATGCAGGAAGCAGAAATGATGTATCAAACTGGAATGAAAATCCTTAATGGAAGCAA CCAAAGCCCTGGAGAGAGTGTCATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC CAGGCAGCGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTAT ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCCACATGGTTTTGGTAAGTAGACTT TAGTGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTTATAGCGGCCACAACTTTT ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA ATTTTTTTTTAAAATGATTAGTTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGC TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA TTTTAAAAATTTTTCTTTAGTTTTTCAAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAAATTCATGGAGTTATTT GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACTCATTTTTAATAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSKTTLTSDESVKDHTTAGRVVAGQIFLD SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG EPCHFPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAAKRRQMQEAEMM YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEK LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMVLVSRL

FIGURE 67

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAG SAPRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVQPEDDADYYCSVGYGFSP

FIGURE 69

GCCGCCCCCCGAGACCGGGCCCGGGGGCGGGGGGGGGATGCGGCCCCGGGGCGG CGATGACCGCGGAGCGCACGCCGGGGCCCGGCCTGACCCCGCCGCCGCCGCTGAGCCC CCCGCCGAGGTCCGGACAGGCCGAGATGACCCCCTGTTGCTGCTGCTGCCGCCGCCGCCGAGGCCCCCCAAAGATGCCGCCGCCGCCGCCGCCGCCGAGGCCCCCAAAGATGGCGGACAGGCCCCCCCAAAGATGCCGACCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGAGTGCCCAGTGGA GGGGGACCCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG AGGACCCCGCCAGCCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC CGGGTGATCGCACGGCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCGTGGCCAGCGGCACCC TCGGCCCGACATCACGTGGATGAAGGACGACCAGGCCTTGACGCCCCAGAGGCCGCTGAGC CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGGAGGACAGCGGCAAATAC ACCTGCCGCGTGTCGAACCGCGGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA GCGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG GGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACTCCACCATCGATGTGGGCGGCCAGAA GTTTGTGGTGCTGCCCACGGGTGACGTGTGGTCGCGGCCCGACGGCTCCTACCTCAATAAGC TGCTCATCACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAACACC GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCCAGGCCCAGTTGCTGGC ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGC**TAG**ACGGCACCGT ${ t ATCTGCAGTGGGCACGGGGGGGGCCGGCCAGACAGGCAGACTGGGAGGAT\overline{ t GGA}GGACGGAGCT}$ GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCCAGGCAGTCTGTGTG TGAGGCATAGCCCCTGGACACACACACACACACACACACTACCTGGATGCATGTATGCAC TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACAGGT GCAGATATGCTGCCTGGACACACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGA TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA CACACATGCACACAGGGCAGATATGCTGCCTGGACACACGCAGACTGCCTTTTGG GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT CCCTGCTCCACCGTCACTCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC CATCCCCGCCTCTGTCCCCTGGCCTTGGCGGCTATTTTTTGCCACCTGCCTTGGGTGCCCAGG AGTCCCCTACTGCTGTGGGGCTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAGCCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCTTGGTA TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTGCAGGGAC

FIGURE 70

MTPSPLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGK
ESLGPDSSSGGQEDPASQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTHSHVEGKV
HOHIHYOC

FIGURE 71

TCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTCCAGTACCTCCAAAGCACCCCAGCCCAG TCCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGCCCTCGAGGAGGCAGCAGCGAGCAGCAGGGGTGGCCAGCG

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGROTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ RPGONCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGOTGVOTRTRICLAEMVSLCS EASEEGOHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTK TPKLLTOTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPY MVMNPETKARRAGOSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG EYFCKAOSDAGAVKSKVAOLIVTASDETPCNPVPESYLIRLPHDCFQNATNSFYYDVGRCPV KTCAGQQDNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRV SAADNGEPMRFGHVYMGNSRVSMTGYKGTFTLHVPODTERLVLTFVDRLQKFVNTTKVLPFN KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKV KASVTFLDPRNISTATAAQTDLNFINDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL DSTOVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENORRNKREDRTFLVGNLEIRERRLFNL DVPESRRCFVKVRAYRSERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLNYRRTDHEDPR VKKTAFOISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYOIEGDRYDYNTVPFN EDDPMSWTEDYLAWWPKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS TRDRDOPNVSAACLEFKCSGMLYDODRVDRTLVKVIPOGSCRRASVNPMLHEYLVNHLPLAV NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT FNCVEROVGROSAFOYLOSTPAOSPAAGTVOGRVPSRROQRASRGGQRQGGVVASLRFPRVA OOPLIN

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC ACCTACCTACCCGTACGCATACATACATATGTGTATATATGTAAACTAGACAAAGATCGCAGATCATAAAGC AAGCTCTGCTTTAGTTTCCAAGAAGATTACAAAGAATTTAGAG ATCCTGTCGATTCATGCCCTTTGGGTTACGGTGTCCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCTGAGCTGATGTTTGATTTT GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA GATGCTTTTCACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTCGCGCTTTTTG CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTGTATGTG TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC CAGTATTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA ACAACGTGCGCTGCCTGTGCCCGGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGC GGGAACCGCCAGCCCCTGGTGTTCTAGGTGTCACCTCCAGCCACACCGGACGGGCCTGTGCCGTGGGGAAGCA CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTC TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA ATTTTTCTTGAACTACTGTAATATGTAGATTTTTTTGTATTATTGCCAATTTGTTATACCAGACAATCTGTTAAT GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATAACAAGTGTAATAAGA TTCCACCAAAGGACATTCTAAATGTTTTCTTGTTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA TTACTGATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCCAGGAAGATGGATTGATGTTCACT AAAAAA

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFTVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGG CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCCTCATGACTT CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC CCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTTTTTGGAATGGAGGAGACGCTCA AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT AGTATACATTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAG AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAAAAAAT TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA ATGGGGTGATGAATGGTCACACCCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCCCTGCCCTGAATCT CATTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT TGGACCAAACCTCG**TGA**GCCAGCCACCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT CCGGGAGCAGTGATGTCAAACTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCGGAAACATCTGCCCATGTGTATTG ATGGCAGAGCTGTTGCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT GCTTCAGAAGTGTTATTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT TTACTTTACATATTCGTTTTCAATACTTGCTGTTCATGTTACACAAGCTTCTTACGGTTTTC TTGTAACAATAATTTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC TAAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG AAAAG

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVKW HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

FIGURE 77

FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG CCTGCCCTTTAAGGGCGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT TCTGTCGCAGGCTGCGAGGAAAGGCCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGCTT CCTCCCGGTCGTCCTCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA TGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC GAGTGTATTATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCC TTCTCCATCATCAGCAATGAGGTGCTGCTCCCTGCCTCGGAACTACATCCAGTGGCT CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCCAACCTGTCCCTCA TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG GGTGTCCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT AGGTATGGTGTGGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT CTGCTCCTGGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGCGGAAGGC GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG TGCCGTCATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCCTCAGTTGTGGGCTTCTATA GCTCTCCACTCTTCCGGAGCCTGCGGCCCAGATGGCACGACACTGCCATGACGCAGATAATT GGGAACTGTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT ${ t CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAG}$ AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG ACCTCAGGACCTGGAATCTGAGAGGGTGGCTGGCAGAGGGGAGCAGAGCCATCTGCACTATT GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA AAAAA

FIGURE 79

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDEDATVNK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGVLGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMPRGMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVGFY
SSPLFRSLRPRWHDTAMTQIIGNCVCLLVLSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY
IVFLYNAAFAGLTTLCLVKTFTAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

FIGURE 80

GGCTGCCGAGGGAAGGCCCCTTGGGTTGGTCTTGGTTGCTTGGCGGCGGCGGTTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

FIGURE 81

FIGURE 82

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCGGTGT GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG GAGTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCACCTTTT GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCA GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG CTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTCACACTTCTTTGGGGATTT TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGAT GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT GCCCTAATGTCAAACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG AGAGACGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TGA**GAGTG TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACC AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA GTTGTTAGCAATTTCATTCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTG TTATTTTTA

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRKQYDTYGEEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPRQQ
DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFIGEGEPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATG**G AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC AGTCTCTGTCGGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATT**TAG**TGCCT ACAGGCCAGCAGCTAGCCATGAAGGCCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTTGCAGTTGCAAACTGTGGCTGGTGAGT GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAGGTTTGGGTTTGAAGCTGAGGAACT ACAAAGTTGATGATTTCTTTTTTTTTTTTTTTTTGCCTGCAATTTTACCTAGCTACCACTAGGTG

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWI DQSLSAAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCC $\verb|CCACCACTGCAGCC| \textbf{ATG} | \texttt{ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA|$ TTTGGAGTGTTTTTCCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTC CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTCAG GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA ACTAAGAAGAAGTCAGCAAGCAAACTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA ACTCTTGAAGGCTATTTGTGTTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA CTGTGGTGCCTGTTTCTTTTCTTTTTTTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT TTTTAGAAGTGTCCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA TGCATGAATTCGATTGGATTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIRRVPVLGSLLNLPGI RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

FIGURE 91

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC ATCAACACCATTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGGCTGGAGCCTGTCCGA TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCCT GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC GGGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAA TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC AAGCTCTACCAGGAGAAGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA GACGCCCATGGTGCCCCCCGGCGGCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGC TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG CTGGCCAGCTTCATCCTCGTCTTTTTTGTGGCCTCCGTGGGAGTTCGATGATGATTGGTGT **GA**CTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACTGGTGGCCTCTGCATATCCT CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA $\verb|CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT| \\$ TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIASGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQHLRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

FIGURE 94

GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT AAACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTTATAAAAC TGTATCAGGTTCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT AGTAAAGGATGTAAACAGATTAAAACGAGAAATTGAGAAAAGGAGGAGCACAGATTCAGG CAGCAAGAGAAGAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACA TGTTTCTAAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC CAGAAACAGATGAAGAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA TTT**TGA**TCCTTTTAACCTTACAAGGAGATTTTTTTTTTTTGGCTGATGGGTAAAGCCAAACAT TTCTATTGTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCAC CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC ATCAGATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCT TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDTQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

FIGURE 96

 $\verb|CCAAGCAGCGCAGCGAACGCCCGCCGCCGCCCACACCCTCTGCGGTCCCCGCGGGGCGCCTGCCACCCTTCCCT| \\$ CCTTCCCCGCGTCCCCGCCTCGCCGGCCAGTCAGCTTGCCGGGGTTCGCCCGCGAAACCCCGAGGTCACCA GGGGACCGTTGCCTGACGCGAGGCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCCTGCTCGCCTCTTCCAC AAGTGCGACGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCCTCCACGAGATCAACGGTGATCAT $\verb|TTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA|$ TGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTG ATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT TTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT GAACCTGGAAGAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAACTCCCAGT ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT AGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCCACAGCCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT GCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACTACTGCTCAAACATCATGAGAGGCTGTTTG GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA GGGTCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG ATAATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCCTCCCAGCTGGACGAATT TCTCGTTCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGC AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGTCCT CCCTTCCGAGCAACGTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG ${\tt AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA}$ GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA GAAAGCCGACAGTGCTGGTGTCCGTCCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGG ${\tt TTATGCAGAGAGAGTGGAGA}{\tt TAA}{\tt TTCTCAAACTCTGAGAAAAAGTGTTCATCAAAAAGTTAAAAGGCACCAGTT$ CACTGGTTTAAGAAGTGCTGACTTTGTTTTCTCATTCAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT CGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAAGAAAAAGCCCAAAAAGC

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLYMQNSELFKDLFVELKRYYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEY
LECVSKYTEQLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGCGPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLF
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

FIGURE 99

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR KFMTVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

FIGURE 100

FIGURE 101

MAVLVLRLTVVLGLLVLFLTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVE FILRSMSRSTGFMEFDDNEGKHSSK

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT ${\tt CAGAGCTGGTCTGCC} {\color{red} {\bf ATG}} {\tt GACATCCTGGTCCCACTCCTGCAGCTGCTGCTGCTTCTTAC}$ CCTGCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCC CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA AGCTGGGCCTTCATGTGGCAGCAAGTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG CTGCCTCACCAGAGAGCCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG AACGACAGCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA AGCCACCACCAGCCTATCTATCTTCCACTGAGAGGGACCTAGCAGAATGAGAGAAGACATT CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT CTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT AACCACG

FIGURE 103

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQI KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV APGEDMRQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFM WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP SSKALICSFPSLQLEQATHQPIYLPLRGT

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAGTTGCCTCATCGCAGG CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA ACTAATATTTATATGACAGAAGAAAAAGATCATCCGTAAAGTAAACATCATCTTGG TCCTGGCTGTTGCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAAA TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCTGTGGTCATCGCTGCATCTG AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTCAGCACAACACTCGCTCCAAT GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAG AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC TTGCCAATTCTGGTTCCCAGCGCAAAGAAGCCCATATACATGGATGATGATGTAATTGTGCA AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG AAGATTGTGATTCAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG CTCATTTAATCCTGGAGTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT GGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTAA AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAG ${\tt ATATACCGAGATCTCAAACATAAAG} {\tt TGA} {\tt AAACAGAATTTGAACTGTAAGCAAGCATTTCTCAG}$ GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA GCAAGCCATGGAAAAAGATGTCAGCTAGGTAAAGATGACAAACTGCCCTGTCTGGCAGTC AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT ACATTTTTC

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR QEEIPVVIAASEDRLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA NLTEWKRQNITNQLEKWMKLNVEEGLYSRTLAGSITTPPLLIVFYQQHSTIDPMWNVRHLGS SAGKRYSPQFVKAAKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

FIGURE 106

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCCGCGCTGCCACG TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT TTCCGATCACTTCAAAGTGGTTAAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT CCGTAAAGTAAACATCATCTTGGTCCTGGGCTGTTGCTCTTCTTACTGGTTTTGCAC CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTCAGGAATTGTAG GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA GGAGATTCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA ACAGCATTCAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG TCAATTTTGACCCTAAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGA CCATATACATGGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA $\tt CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT$ CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTTGCAAAC CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGT AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCG ${\tt TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT}$ GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA $\tt CTGATATGGCTAGCTAGCTGGTACAGATAATTCAAAACTGCTGTTGGTTTTAATTTT$ AAAAA

FIGURE 108

FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH PRSPAMKAATCCSPEGPWPSLEPRT

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA GTTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAA TC**ATG**TCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCATTTGGTTAT TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCTCTGGGTGGCTGTGCTG AGCCTGGGAACTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCT ${ t TTCGGGCATTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA}$ TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA TACTACAACTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT TGTCCAAGAACTCAAGTCACTTTACATCTATTAACTGCTTTGGAGACTTCATAATTTTTCTA GGAAAGGTGTTAGTGGTGTTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG ATAGTTTTTTTTTTTGTGTTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTTGAT CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT CGTAAAAAGGAGCAACAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA GGAAAACATTTCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT AGTGAATTTTTTTTAAAAGACCTAATAAACCCTATTCTTCCTCAAAA

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVILGLLFVCGVLWWLYYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

FIGURE 112

TGTGGTGAAAATTTTTTGAAAAAAAAATTGCCTTCTTCAAACAAGGGTGTCATTCTGATATT TATGAGGACTGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCCTTGTTTTGCTGG TGACTGGAGTACATTCAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCCAAGTTCACTGTG CCTCAGATCAACTGCGATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATG TCCAGCAGGATGCCAAGACCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACT CCAGTGTGTGTGCGCTGCCGTACACAGTGGTGTTGATAATTCAGGAGGGAAAATACTT GTTCGGAAGGTTGCTGGACAGTCTGGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTT ATCCCTACCACGATGGAGAGAATCCTTTATCGTCTTAGAAAGTAAACCCAAAAAAGGGTGTAA CCTACCCATCAGCTCTTACATACTCATCATCGAAAAGTCCAGCTGCCCAAGCAGGTGAGACC ACAAAAGCCTATCAGAGGCCACCTATTCCAGGGACAACTGCACAGCCGGTCACTCTGATGCA TGCAAAATTGACTTGTCGTTTTTAATTGATGGGAGCACCAGCATTGGCAAACGGCGATTCCG AGGTCGGGCCATCTCCTTTGTGACCAAGAACTTCTTTTCCAAAGCCAATGGAAACAGAAGCG GGGCTCCCAATGTGGTGGTGGTGATGGTGGATGGCTGGCCCACGGACAAAGTGGAGGAGGCT TTGTGACCAACCTCACCAAAGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTG CAGTACACCTACGAACAGCGGCTGGAGTTTGGGTTCGACAAGTACAGCAGCAAGCCTGACAT CCTCAACGCCATCAAGAGGGTGGGCTACTGGAGTGGTGGCACCAGCACGGGGGCTGCCATCA ACTTCGCCCTGGAGCAGCTCTTCAAGAAGTCCAAGCCCAACAAGAGGAAGTTAATGATCCTC ATCACCGACGGAGGTCCTACGACGACGTCCGGATCCCAGCCATGGCTGCCCATCTGAAGGG AGTGATCACCTATGCGATAGGCGTTGCCTGGGCTGCCCAAGAGGGAGCTAGAAGTCATTGCCA CTCACCCCGCCAGAGACCACTCCTTCTTTGTGGACGAGTTTGACAACCTCCATCAGTATGTC ATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGCTGGAGATTTTACATTTTGACAATT GTTTTCAAAATAAATGTTCGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAGAGCTTTT GTGAGATTTTTAAGTTGTTATTTCTGATTTGAACTCTGTAACCCTCAGCAAGTTTCATTTTT

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQ
RQDPSGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLQPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDI
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCGGGGGGCGCTGAGAGGACACGAGCTCT**A** TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACTCC AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGC GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCCTAGAGTATAACAAAG CCATCCGGAACTACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT GTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT CCTAGAACTCGGAAGAGGTGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG GTCCACCTTCGACGCGGTGATCACCCCCTATGGGGGGTGCATCCTGGGGGCTGGGGGGTACA TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCAAACGGAGCAGGTC GAAATTTCAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGGGAAGCCTGCCAAACAGAAGGTCCCA CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCCT AATTGCTTTTGGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLWVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

FIGURE 116

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA A**ATG**CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA AACAGTGTACTATTCTGTCGAATACCAGGGGGGGTACGAGGCCTGTACACGAGCCACATCT GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG GAGGAGGTGGATGCCTGTGCCACGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT CTCA**TAG**GTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAAGTAGGAAGA GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTTGTTGTGCTCCTTTTTTC TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA AAAAAAAAA

FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW
SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

FIGURE 119

CGGACGCTGGGCCGCCACCTCCGGAACAAGCCATGCTGGCGGCGACGGTGGCAGCGGCGTG GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGACTTCTACGACTTCAAGGCGG TCAACATCCGGGGCAAACTGGTGTCGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG AGACCTGGGCCCCCACCACTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCGCCGCACCTACAGTGTCTCATTCCCC ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA GACTTCTGGGAAGGAGCCCACCTGGAACTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG TGGTAGGGGCTTGGGACCCAACTGTGTCAGTGGAGGAGGTCAGACCCCAGATCACAGCGCTC GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA**TAA**CCACCGCGTCTCCTCCTCCACCA CCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAACTCAAATGGTGCTTCAAAGGGAG AGACCCACTGACTCTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT GTTATTTCCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA AACAATACCTCACGATATAAAATAAAAATGAAAGTATCCTCCTCAAAAA

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ HYRALQQLQRDLGPHHFNVLAFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG AHPAFKYLAQTSGKEPTWNFWKYLVAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

FIGURE 121

CGGACGCGTGGGCGGGCCGGGACGCCAAAGCGAGCCATGCTCTACGTCGGGATGC TGCGCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGGCCCGGCCCCTCTCT CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAGGGGTGCACCAAAAAGCATCTTA ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC TTGGTCGTCCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC CTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG TCTGTGAACCCAGCCTACCAGGCTATGGAACTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT GTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCTTGAAGAGTCAGAGGCTCCCAGATCTGACC ACAGTCATCTCGGTGGATGCCCCTTTGCCGGGGACCCTGCTCCTGGATGAAGTGGTGGCGGC TGGCAGCACGCCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCCTGTCCTGCCATG ACCCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCACCCTCTCC CACTACAACATTGTCAACAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAG GCACAATGATGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG GTGTCATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGG CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGC ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG CTCGAGGACTTCTTTCACACACCCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTCGGCTGAAGGACGGGGAGGAGACCACGG TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC GTGTTTGTCACAAACTACCCCCTCACCATTTCAGGAAAGATCCAGAAATTCAAACTTCGAGA $\mathsf{GCAGATGGAACGACATCTAAATCTG}$ TGA $\mathsf{ATAAAGCAGGCCTGTCCTGGCCGGTTGGCTT$ GACTCTCTCCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC TCCATCCCCCACATTCCCCTGTCTGTCCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT GAAAAAAAAAAAAAA

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDFFHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277, 314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

FIGURE 124

GAGCAGGACGGAGCC**ATG**GACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC AGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC GTGGACGTCTGCACCGAGGCCGTGGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC AGTGCGGGGTTGCGGTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC TTCTGGCGTTCATCCAGCTGCAGCATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCT GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA GCTAATGTGACTGTGCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACT CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG ACCCACATCCACCACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG $\texttt{TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCCTACTG} \underline{\textbf{T}}$ **GA**GCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT CATCACTTCCTGTTCCCACCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTTGGGAAATAAAATACCGTTGTATAT ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC TCCGCTTGTCCTCTTGTGATGTTAGGACAGAGTGAGAGAGTCAGCTGTCACGGGGAAGGTG GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC CCAATTCGCCCTATAGTGAGTCGTA

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPPLVRLPPPEPTT VASTTSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVEHEASRDEEPRLTGGAAGHQDRSNSG QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

FIGURE 126

CGGGACTCGGCGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCCATGGAGT TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCT**TAA**GAAAACCGGCCACTTC AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCCTATCCCCTCTA ACACCATTCCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCCT CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCGTGGCCCTCCATCAC CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCAC TTTGTCATCATTCTTCATGGACTCCTTTCACTCCTTTAACAAAAACCTTGCTTCCTTATCCC ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG AGGAGTCCCCATCTGCCCGCCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCT ACTCTGCCCCTGGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG GTTGGGGCCAGCACCGGGATGGATGGAGGGGGGGGGGGCCTTTGCTTCTCTGCCTACG TCCCCTTAGATGGGCAGCAGGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA GCGGTGAGCGAGGTTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG GTTGAAGGTCATAACGAGAGTGGGAACTCAACCCAGATCCCGCCCCTCCTGTCCTCTGTGTT CCCGCGGAAACCAACCAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT CCTCAACAACAACAGAAAAAAGGAATAAAATATCCTTTGTTTCCT

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCK CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLEPLDANVDA

FIGURE 128

FIGURE 129

MKIPVLPAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSAFKADE FLNWHALFESIKRKLPFLNWDAFPKLKGLRSATPDAQ

FIGURE 130

FIGURE 131

 ${\tt MGVEIAFASVILTCLSLLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS}\ .$

FIGURE 132

GGGGAATCTGCAGTAGGTCTGCCGGCG**ATG**GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG GCTGCTGTTGTTCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC TGCAGCTGCTACCATGGTGTCATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG GAAGATGATGCCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTTGAGCACTTTATT TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACTTGTTGATGCAGAATACACCAAAAAC CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT TCCCAAAATGTTGAAAACTGAACTA**TAG**TAGTCATCATAGGACCATAGTCCTCTTTGTGGCA ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA TCTGCTATCAAGCCAAATACCTGGTTTTCCTTATCATGCTGCACCCAGAGCAACTCTTGAGA AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAAAACCATAAACTCTGTTACTCAG GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

FIGURE 133

MEWWASSPLRLWLLLFLLPSAQGRQKESGSKWKVFIDQINRSLENYEPCSSQNCSCYHGVIE EDLTPFRGGISRKMMAEVVRRKLGTHYQITKNRLYRENDCMFPSRCSGVEHFILEVIGRLPD MEMVINVRDYPQVPKWMEPAIPVFSFSKTSEYHDIMYPAWTFWEGGPAVWPIYPTGLGRWDL FREDLVRSAAQWPWKKKNSTAYFRGSRTSPERDPLILLSRKNPKLVDAEYTKNQAWKSMKDT LGKPAAKDVHLVDHCKYKYLFNFRGVAASFRFKHLFLCGSLVFHVGDEWLEFFYPQLKPWVH YIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMDDITCYWENLLSEYSKFLSY NVTRRKGYDQIIPKMLKTEL

FIGURE 134

CACCCCTCCATTTCTCGCCATGCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTTTCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT $\tt CCGCAGCATCCTTGCCCCCCTGGCATGGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC$ ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCAGCTGGTGATGCGGTACTGGGAGCC CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCCGAGCC TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCCACCTGCGCCACCCAGTGTGTG TGGAGCTGCTGACAGTGCTGTGGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT TTCCTCCTTACCCTCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT $\texttt{CCGGGCCCAGCTACAAAGAAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAG} \textbf{\underline{T}}$ **GA**GGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCTCTCCCACTGAATTCTAAATCCTTAAC ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA $\tt CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGGTATGGAAGTCATCTGTCCC$ TTCACTGTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC $\tt CTGACCACTCCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT$ GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA GGCCCCAACCTTGCCTCACCACTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC

FIGURE 135

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP LAWDLGLLLLFVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP RALRLFSHLRHPVCVELLTVLWVVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

FIGURE 136

 $\tt CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA$ AGAAATTGCCAAACCATGTCTTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTCATCTGGAATT TCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAAGAACCTTCGGCTCGCGTGCTTCTG AGCTGCTGTGGATGCCTCCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCACTGAGATCC CTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT TCCCCACTACAATGTGATAGAACGCGTGAACTGGATGTACTTCTATGAGTATGAGCCGATTT ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATCCATTT CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCCAGGCCATTAGAGTTAC TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTCTTATTAGGCCAAG AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGC TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTT GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTGGCAGGTCATGCTAAGGAACACC ACATGCCATTAT**TAA**CTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG AACTGAAACTCATGAAAAACCCAGACTGGAGGACTGGAGGGTTACACTTGTGATTTATTAGTC AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTCACAGTTATTATTATTTAAAAATTA CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAAT ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQD FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEK EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD LVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG FSSKEIITFWQVMLRNTTCHY

FIGURE 138

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN PNKVDDLSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

FIGURE 140

CATTTCTGAAACTAATCGTGTCAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT CAAGCTTATAGTTGAAATATTTTTCAGGAATTAC**ATG**AATGACAGTCTTCGAACCAATGTGT TTGTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT CAGATTCCGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAAT CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAAACTATGAATTAC TGGAAAAGAAGTAGAAAAAGAAAGTAGCCTTACAAGAAGCCAAATTAAAAGCAAAGGGA TTGAATCCGGATGGAACTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT GGTTCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGACATGG TCATAAAAGGAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT **CTGA**CTTTCTCTTTCGAGCCTGCATCAGTTCTTGGTTTTTGCCTATCTACAGTGTGATGT ATGGACTCAATCAAAAACATTAAACGCAAACTGATTAGGATTTGATTTCTTGAAACCCTCTA GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACATT AAAATGCCCTAGCAGTATCTAATTAAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTACAAGGAAATAAAATACAAAT CTTGTTTTTTCTAAAAAAAAAAAAAAAAAAAAAGT

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ
SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSGHGRHRR

FIGURE 142

FIGURE 143

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTCATGGC GGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCCGATACAAGCACCCG TCTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCCCAGAGGCAGGAAAGAGCGGTG GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG AGACCTGCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACAT GCTGGGACCAGCCAAGGAGACTAACATTGCTGTTCTGGTGCCTGCTCACGGTGACCTTCT CCATCAAGATGTTCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGCCATGCTGGTGCAAGTGGTGCG GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG GGACTGGCAGTGGTGGGTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC CCAGACCCACCGGGACGCACTGACCATGTCGGAGGACAGACCCATGCTGCAGTTCCTCCTGC ACACCAGCTTCCTGTCTCCCCTGTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCCTCAACTGCACACTTCTGCTCAAGACGC TGGGAGGCTATTCCTGGGGCCTGGGCCCAGCTCCTTACTATCCCCCGACCCATCCTCAGCC AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT GGACGGCTGCCTGCCAGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA ${\tt GGCTCC} \underline{\textbf{TAG}} {\tt CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCCTGGGGCAGCGGGACA}$ CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCTTGGACGCCGAGTTTCTGCCTCA GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC GATTTTTAA

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLEPGLASMTQNLEPLLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFLTFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVSLQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAARI
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLAGS

FIGURE 145

FIGURE 146

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTT AATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGGCATGGCAAGGTTTGCTTAAAGGAGC CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG TATTCAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG $\tt GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA$ AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG GCAGGGAAAATTCTGAAAACACCACTGCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACTACGCTGTGCGTCTCCTGCGG CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCC GGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAA TAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA CATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCCAGAAAG TGCGGCTCATCTGATTCAGGCCAGTGAAAGACGTGTTCACCTCGTCGTGTCCCGCCAGGTTCGGCAGCGGAGCC GAGGAGTCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTTTGTTGAATGTGGATGGGGTCGAACTGACA GAGGTCAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT GTGACTGGTCCCCATCCTGGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA $\tt TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG$ CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGA ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTA<u>TAG</u>AATCAATGATGGGTCAGAGGAAAACAGAAAAA AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAAATTA AAACTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTTTCTATTCAAT AAAAAGCCCTAAAACCAACTAAAATGATTGATTTGTATACCCCACTGAATTCAAGCTGATTTAAATTTAAAATTT GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTAATTTACAGCTAAAATATTTTTTAAAATGCA TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTCAGAAGTTAAA

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SESLSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVVSRQVRQRS
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

FIGURE 148

FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT IFDYKHGYIASRVLSRRACFILKMDHQNIPPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

FIGURE 150

ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGCCAGCAGCTTCTCCAAGGCACG GGAGGAAGAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA GCCCCATCCTTGCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCATGTAGAATGAACC GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTCGTATTTGGA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPPITYSL CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK PVSELRANFTLQDRGAGPRVEMICQASSGSPPITNSLIGKDGQVHLQQRPCHRQPANFSFLP SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

FIGURE 152

 ${\tt GGTCCTTA} \underline{\textbf{ATG}} {\tt GCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCGCTTCTGCTCCTG}$ CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCTGC AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC AGCCACCACCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAAGGCTCCTGTGAGCACG GTCTTGATCAAACTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTTGGTGCTACC GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGGATTCTTTCCGTGTCCTGAAAGAG AATTTTTAAATTATTTAATAAGAAAAATTTATTATTAATGATTGTTTCCTTTAGTAATTAT

FIGURE 153

MAAAAATKILLCLPLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL HYDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHELWLLIKREFGFYSKSQYRTWQKKLA EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACCATCAAG CTCTTGTGGCAGGTAACTGTGCACCACCACCTGGAATGCCATCCTGCTCCCGTTCGTCTA CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC AGAACTGCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCCTGCAGT TGGGCAGGACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC ACCCTGGAGCTGTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC CAAGCTGCGGGAGCTCTGGCTTCGCAACACCCCATCGAAAGCATCCCCTCTTACGCCTTCA ACCGGGTGCCCTCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA AGACATGCCCAATCTCACCCCCTGGTGGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCATG CAACTTGGCCCACAATAACCTCTTCTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC TGGTGGAGTTGCATCTACACCACACCCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCATGCTCCCAT GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT CGGACTCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC CTCCCGCCACCCAAGGATCTCTGTCCTCAACGACGCACCTTGAACTTTTCCCACGTGCTGC TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACTACAGCTTCTTCACCACAGT AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCCTA CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC CGTGTGCCCAAGCAGTGCCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT GGATGAAGTCATGAAGACCACCAAGATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC AGCAGCAACAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTC ATGACCATATTAACTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC CTGGGGAACTCTCTGCACCCCACAGTCACCACTATCTCTGAACCTTATATAATTCAGACCCA TACCAAGGACAAGGTACAGGAAACTCAAATA**TGA**CTCCCCTCCCCCAAAAAACTTATAAAAT GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAAATTTAAAGA CAAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEMSGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHNPWNCDCDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

 ${\tt CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCC}$ TGTCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGGAAAGAGGGGAAGTCGTGGGTTATACCAT GCAAGAGCTGCCGAAATGGCTCATGGGGGGGGTACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAG TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT GTTGGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTCATGCTAAACCTGGGTTTGTCATCCAACTAA GATTTGTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTCGTGATGGAGACAAC CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGCCAGCTCCTATCCAGAGCATAGGATCCTCACT CCACGTCCTCTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT GCTCCTCATCCCCTTGTTTCCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG TGGGTACCAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTT TCTTTTGTAACAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGG AAACAGCCCATCTGCATAAAAGCCTGCCGAGAACCAAAGATTTCAGACCTGGTGAGAAGGAGTTCTTCCGAT CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAG TATGAGTGCATCTCACCCTTCTACCGCCGCCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG TGGGCGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTTGC GCTGGCCGTGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGCAGCCTACACAAGGGAGCGTGG TTCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGG GAAGGTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATG AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCCTGCTTGATGCT GACATCGCCATCCTGAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAG GCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGGTGGTCAGTGTGGGGACTCGCTGCTGTGTGAGGAGCAG CATGAGGACCATGGCATCCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCCTTC TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCCTTCCCGGGACGAGCATCTCCTGAGCCACGCT GGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTCACCAAG GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAA<mark>TGA</mark>ACCATGCTCATGCACTCCTTGAGAAGTGTTTC TGTATATCCGTCTGTACGTGTCATTGCGTGAAGCAGTGTGGGCCTGAAGTGTGATTTGGCCTGTGAACTTGG CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT ATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTG ${\tt ACCAGGGAAGATCTGGGCTTCATGAGGCCCCTTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC}$

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSG
VVSVVDSLLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAAC**ATG**GGC TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCTAAAG CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAAACTCTGACT AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAG AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC CTCGTTCCCCACCGGAACAGAGAGAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT CCTGCAGAGGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT TTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG AATGAAAATTTCCCGGCCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCACGAGTCTGG AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA TATCAACATCACAGTGGATTTCTGGTTTGGTGCA**TGA**CCCTGGATCTTTTGGTGATGTTTGG AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTTGTATTTTCTTAGCAGAGCT CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCT GTGAAGTGGTGTCAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA GGACACAGTGAACTTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTCGGCTGCAAAGGCAG CAGTAGCTGAGCTGCTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT TCCAGTGATGCCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTTGTAAAATGA TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 161

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

FIGURE 162

CGTGGGCCGGGGTCGCGCGGGGCTGTGGGCGCCCGGAGGAGCGACCGCCGCAGTTCTC GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCA**ATG** CCCGGCAGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA CGGGAGCGTGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG CCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTCGGCCACGTGCCCGGGGAATTCCCGG TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG GTCCTCCCCATCACAGAGTTCCTCGTGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC GCAACTTCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTTGGGG CCCACCTAATTCAGACCTTCCAAAAGATGACCGTGACCTTGAACTTCCTGGGGAGCCCTCC TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCCTGGGGACTAC TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG AACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT GGAGACTCCATCTGAGTACCTGGAAATTGTTCGTGAGAACCACGGGCTGCTCCCGCCCCTCT ATAAGTCTGTCAAAACTTACACCGTG**TGA**GCACTCCCCCTCCCCACCCCATCTCAGTGTTAA CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCATT TGCGTGGGGCTGTTGGCCTGGATCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC CCTCCCTCTCTGTCACCCCTGACCCCAGCCATTCACCCATCTGTACAGTCCAGCCACTGACA TAAGCCCCACTCGGTTACCACCCCCTTGACCCCCTACCTTTGAAGAGGCTTCGTGCAGGACT TTGATGCTTGGGGTGTTCCGTGTTGACTCCTAGGTGGGCCTGGCCCACTGCCCATTCCT CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCCTCCCCCAGACAGCCCTAC CTGTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC CACACACACAGAAATATAAACACATGCGTCACATGGGCATTTCAGATGATCAGCTCTGTA TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA AGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGCCTGGATGGGGGGCAGGACT AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA AGCCGGGCGTGGTGGTGGCTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG GTGCGAACCCGGGAGGCGGAGCTTGCAGTGAGCCCAGATGGCGCCACTGCACTCCAGCCTGA GTGACAGAGCGAGACTCTGTCTCCA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 164

FIGURE 165

 $\verb|MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTH| \\ FPICIFCCGCCHRSKCGMCCKT$

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC CTGGATCTTCCACC**ATG**TTCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC AGTCTCCTTTGGTATCCGCAAACTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGGGCTA CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC GGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTCGTCGAAGTGG TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC GGTCCTGTGGGGGTTAGGAGTGCTGATTCGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTTGCCAAAT AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTCAGAGAGCCATGGTGAAGGCCTG CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC CATTGCCAGGCAGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC AAGGACAGGAGCCGCTCC**TGA**GCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCACTGCTGTCCTTTCCAGACTCCAGGG CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACTCCCCA CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGGCGGCCACCCG CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTTCGGCCTCAACATCGCCCCCAGC CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAAATG GGCCGCTGACTGGGCCATGGGGAGAACGTGTTCGTACTCCAGGCTAACCCTGAACTCCCC ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG AGGACACATCACGTTCAGTGTTTCAAGTACAGGCCCACAAAACGGGGCACGGCAGGCCTGAG TGA

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSGSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLLPLRIALAFTG
ISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRPRNGGICVANH
TSPIDVIILASDGYYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLPPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQQKLYSKMIVGNHKDRSRS

FIGURE 168

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA CGTCTCCTCCAGGGATGCCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCTGGGCCTGGCTCCAGACACCTTTGA CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG AAATGGCCCACCATGCCTGCTGCGGGAATCCTGGGAGGCCAGGAGACCTGGGAGGAC AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG GCCCTGCAGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGTTTCCG AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCCACAGATTTGGGGAGAAGAGGCGGGGCTGT GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCCTG GAAGACTCTGCTCTTGGCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCC**TGA**AAGTCCA ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL YMRHFPFKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

FIGURE 170

GTGGCTTCATTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCA CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG GTCGGTTCCGTTGGTGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC TATTGTCTGGACCTTCAACACCCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCA TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG CTCAGCAAACTGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC CCATAATGGGTCCATCCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT GAAGGTGCTGATGACCCAGATTCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCT AGTACATTGAAGAGAAGAGAGAGTGGACATTTGTCGGGAAACTCCTAACATATGCCCCCAT TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC TGCTCACGATGCCAGACACCAAGGCTATTTGCCTATGAGAATGTTATC**TAG**ACAGCAGTG CACTCCCCTAAGTCTCTGCTCA

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

FIGURE 172

CTGGTTCCCCAACATGCCTCACCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC ${\tt TCTGGACCGTGAAAGAGCTGGTCGGTTCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC}$ CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACACCCCTCTTGTCACCATAC AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA GATGGAGGCTACTCCCTGAAGCTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGT GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAAACTTCTCAAGCCCC ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCTCCATGGTCCTCCT GTGTCTCCTGTTGGTGCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATGAAGATGCTGCTC
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCATCACTGAGAAGTCCAATTCCAGTCTATCAACATTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTTGACTTAA

FIGURE 174

MKMLLLCLGLTLVCVHAEEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ IHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

FIGURE 175

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGATGAGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAAGTCCTGGTGGAAATAGTGAA
AAAATGTGGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV KHCTDQISFKKRLSLKKSWWK

FIGURE 179

FIGURE 180

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK YKSSQKQHSPVPEKAIPLITPGSATTC

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCC AGTGTTTCCCACAGCCCCCAAAACGGAACTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT ACAGAGGAGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCTGAAGCCCTGG CTGGGAGAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC GCCCGCCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGCCGCCC TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG CACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG ATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC CTGGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG ATGCTGCACCCAGGACATTGTTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTC CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG CGTTGATGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAA TTGATCATGCGCGCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA G**TGA**CTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEGILLSGGDKWSRHRRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFFKDKAKSKTLDFIDVLLLSKDEDG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

FIGURE 183

FIGURE 184

 $\verb|MYKLASCCLLFTGFLNPLLSLPLLDSREISFQLSAPHEDARLTPEELERASLLQILPEMLGA| \\ ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV| \\$

FIGURE 185

FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

FIGURE 187

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT**ATG**AACACGTGGCTGCTGT TCCTCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG CTCCTGCTGGACTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT GAGTATGTCCCCCACCTAAGCCCCCGATCCCCCCAAGGCTGGTGGTCAGAGCTGCTCATC TTACACCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT GTCCCCGTGTGCGCATGTGTTCAGGGTCAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAAC TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCT**TGA**AGTAACAAGTTTAAAATGTTCAGA GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT AACAACACCTCCCTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCG TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC ACGGGCAGCAGAGTGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTCAGGCTCTGGGCTCACCTCCATCTCCAGA GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCCACCAAGA GTTTTAAAACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA ATTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDFLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ AQQEAELTPRPAGVVPGA

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAG ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACAACTGGGAGACTGGGGATGA CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAG GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACTGAGCGCCTTTGCTG CTGTTTCCTCTGTCAGGTCTCCTGGGGATGGTGGCCCACATGATGTATTCACAAGTC TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAAC CCGAACTGCCTACCACCATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG TCTCTTGAGCATGGTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC CTAAGGGATTCCTGGGTGCCACTGCTCTTTTCCTCTACAGCTCCATCTTGTTTCACCCAC CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP VSLDGDTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLETTCLE LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTCAG GACATTCGCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

FIGURE 192

MWLPLGLLSLCLSPLPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQGSGME HRNHLCFCDLYDRATSPPLKCSLL

FIGURE 193

CCGCCATGGCAGGCATCAAAGCTTTGATTAGTTTGTCCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTT ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTC TTACAACGGGCATTGTCGTGTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA GCTTGTGCACTTGTTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG CAATGACGACTTCAGCTGGCAGCAGTGG**TGA**AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTT GTTGGCCATTCACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT CTTTTGGAAAAGCTTGACTGATTTCACACTTATCTATAGTATGCTTTTTGTGGTGTCCTGCTGAATTTAAATAT TTATGTGTTTTCCTGTTAGGTTGATTTTTTTTGGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA TTTGCATTGGTTAGGAATTCAGAATTCCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAAATTATT CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTTGTGGATTG AAAATTATTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA GGGAAGAATGACATTGAAATTCCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTTGTGATCTCCTATC AACCTTTCATGTTTTACCCTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC ATCATATATGCCAGAAAACCTTCCTCTGCTTCCTCCTTTTGACTTATTTGGTATGTTGTATATATTACATAAAA CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACTTGTTAGTCTTACAGATAATTCATGCA TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAAG TATTTTTAAGACAAGTTTCCTGTATACCTCTGAACTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT ACCTGACCAAAAATTCCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA TTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTCAGCTTTTTTACTAAAGATGCCTAA AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTCAGATTTCAAGAGGAA GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACTCTTTGTGCTTGTGATCTACTGGACTTT TTTTTTGCAGGAAGTGCATTCTCTGGTCCTTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCACTGCTACTG TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT GTGTCTTTGACCTTGTATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT AGATTTTAAATATCTATTTTAAAAAAAAAA

FIGURE 194

 $\label{thm:calpivnkywplfvlffyilspipyciarrlvddtdam snackelaiflttgivvsafglpivfarahliewgacalvltgntvifatilgfflvfgsnddfswqqw$

FIGURE 195

CTGAAATATATTTTGAGATAGGTTTAGAATGTCA

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSVDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLTSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDAVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVIINEYGSCTCHQQPARECEV

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCCGGGTCCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCGCGCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGTAG
CCCCGTGCGTGTCCGCAGCATCTCGGCTCTGGGCCGGACGACGACGAGGACCTTGG
CTGTTTTCCTGGCGTCCCGCGCGGGCCCCTACGCTTCCACGGGCCGGGCGCGCTTCAAGCCCGCG
GGCCCCGAGGACTGCCCCGCGCGGGCCCCTTCCGCGCAACGCGGAGGCCCAGCCGTG
GGCCCCGAGGACTGCCGCGGACCCCTTCGGCCTTCCGCGCAACGCGGAGGCCCAGCCGTG
GGCCCCGAGGACTGCGCGCCCCCCT

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR AGRLRFHGPGALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 199

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP FARDAVKKCFAVCLA

FIGURE 201

CCGAGGGAGTCTCCTCCAGACCCTCCCTTCCCGTTGCTCCAAACTAATACGGACTGAACGGATCGCTGCGAGGGT AACTGATCAAGTACTTTGAAA**ATG**ACTTCGAAATTTATCTTGGTGTCCTTCATACTTGCTGCACTGAGTCTTTC AACCACCTTTTCTCTCCAACTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT TATATAAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT TTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTGCAGAGAATCATGGGATTGTTGC AAATGATATGTTTGATCCTATTCGGAACAAATCTTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT GGGAAGAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA ACAGATGTAAAAATACATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG AGTTGCCAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG ATGACATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTCAGATATTGACAAGAAGTTA GGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATACCCTGATTG ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT CATCCTAATCTTACTGTTTACAAAAAAAAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCA ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC TCAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA CAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA CTATACTCCTCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCATACCCTTATTTCATAGGGGTC TCTCTTGGCAGCATTATAGTGATTGTATTTTTTGTAATTTTCATTAAGCATTTAATTCACAGTCAAATACCTGC CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCC**TAA**TGTTACTTTGAAGTGGATTTGCATA TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCTGGGAAACCAGTTCCAAACATCTGC ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATTTTAGCAACTTTGCACTATGTAAAGTACCTTATAT ATTGCACTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC TTGATTGAAAATGACAACTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACTGAAGGAAATT TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGA AAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT ATTCGTTCTAAATATATTGTTTCTGTCATAAAATTATTGTGATTTCCTGATGAGTCATATTACTGTGATTTCA TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA AAATCAAATTGGATAAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTCGCTCCCACGGGCGGGACCTTTGTAACTGCGGGAGGCCCAG AGAGAGGCCAAGCCCCTTGCCTTGGGTCACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA ${\tt CCAGATCCAGAGGCAACAGGGAC}$ AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA AGCAGCCACCACCACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCTGACGTTGCC $\tt CCTGCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAAACTGTT$ CAGCTCCCACAGGTTTCAGGTCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC CTGCTGATTCTGCTCCGGCTGTGGCCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCTGGACTGATGAGTTTGCTGTATC AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT CAGGCTGGCATGTTCACTGGGCTGGTGTTACGACAGAGAACCTGACAGTCACTGGCCAGTTA TCACTTCAGATTACAAATCACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEQPPPTPV SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELILDL KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLSSFTTSLRSWMPVVVVVSFILDI VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLLRLKQMNVQLAAKIQHLEFS CSEKPLD

FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCGCCAACAT **G**CTCTGTCTGTGCCTGTACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTTTG ÄGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC CAGGAATTCTCCACCTACCGCCAGTGGAAGCAGAAAATTGTACAAGCTGGAGATAAGGACCT TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG CATGGATAAAAACGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC ACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG GGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGGCAGACGGGGATGTGGTG GAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCCTGG ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT GGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA GACCTTGCAGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTGAACAGCGC GGACCCCGGCGTGTTTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG CCAGCTACCCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAGGGGGGCCTTCGGGCTGTACAGGGGGGCCTGGCCCCCAACTTCATGAAGGTCATCCCAGCTGTGAGCATCAGCTACGTGG ATCTCCCTGTGCCCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAG TGGGAGTGCAGGGGCTCGGGCTGCCTGGCCTGCACAGAAGGCAAGTGCTGGGGCTCA TGGTGCTCTGAGCTGGCCTGGACCCTGTCAGGATGGGCCCCACCTCAGAACCAAACTCACTG TCCCACTGTGGCATGAGGCAGTGGAGCACCATGTTTGAGGGCGAAGGGCAGAGCGTTTGT GTGTTCTGGGGAGGAAGGAAAAGGTGTTGGAGGCCTTAATTATGGACTGTTGGGAAAAGGG TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCCTGTCCAACC CCAGCAGGGGCGCAGCCAGCCCCACATTCCACTTGTCACTGCTTGGAACCTATTT AAGCCCGCCCAGTGGGATGGGAGGAGGAGGAGGAGGGGGGCCTTGGGCCGCTGCAGTCACAT CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTG GCTCCTTTCCTTTGGCAGGTTGGGGAAGGGCTTGCCCCCAGCCTTAGGATTTCAGGGTTTGA CTGGGGGCGTGGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTCCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG
LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 207

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCAT CAATTGCACTCATCATTGGCTTTGGTATTTCAGGGAGACACTCCATCACAGTCACTGTC GCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAA ACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCA AAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTTT GCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGC TGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATA AAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTG CGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCA GGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCA TGAAGGTTGTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAA AATGACATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAAGGCGGAG GGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAATGTGCCTTGGCCACAAAAAAG CATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTCACCACCAGATATGACCTAG TTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCA GACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAG GGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTG TAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCA CAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGG GGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCT TGGCTTCTCTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAA ACAGAGCAGTCGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAA

FIGURE 208

MASLGQILFWSIISIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD QGANFSEVSNTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRR SHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

FIGURE 209

 ${\tt GAATTTGTAGAAGACAGCGGCGTTGCC} \underline{\textbf{ATG}} {\tt GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT}$ GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCTATCATCTTCA CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG GTTTGGGAT**TGA**AGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG AGTTTTGTAACCTTCAAGTGCTGTTCAGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA AAATCCTCACTGCCAGCCCCTCTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCACCTGAC

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLLQEMKTLFLNTEYLMPFL LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS RHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTCGAAAAGATTCCGCAATAAAACT TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT $\tt TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACC \textbf{ATG} AAGGGCATCCTCG$ TTGCTGGTATCACTGCAGTGCTTGTTGCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT GCTGAAGAACACTTTCATTTTGTAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG AATCTAATGGAACTTCCTGTCGTGGGAAGCCCTGGAAATGCTATGAAGAACAGTGTGTC TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCA TCCCACACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG ${\tt GGGACTGCTGCCC} {\tt TGA} {\tt GGTCCTGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTCTG}$ AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTTCCCTGCTCTGCCCCGTTTAACTGC CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK TLGGVIFRKFECANVNSLTPTSAPTTSHNVGSKASLYLLALASLLLRGLLP

FIGURE 213

ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA ATGATGCCATGGCCGACGCCGACCGCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA GGGGGTTGTGATTAATGCCGGAAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG ACTTCAAGCCTGCCGCGCTCCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCCTCACCCACAG CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCCT TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGCACCTATCAACAATGTC CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG AAGCAGTCCCAGCCTGCCAGCCCTGCCCAGCCCTGGCTTTTTTGGAAACGGGTCA GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTCACAGAGATGCAACCAATA GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACTGGAAAA CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP SPSPTAMPSPEDLRLVLMPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC TYQQCPCNRLREECPLDTSLCTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

FIGURE 215

GCGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAGTGCGAAGAGCA GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTCACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCT ${\tt CAGTTCCATGGCAAGTGGCCCTTCTCCCGGTTCCTGTTCTTCAAGAGCCGGCATCGGCCTGGCCTCGTTTCT}$ CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCCATGTACCACA CCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC CGTGGGGCTGCAGCACCCAGCTGTGGTCAGTGCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGCACGTCT $\verb|CCTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTG|\\$ GTGTGGTGGCTGGCCTGTGGAACCAGCGGCGGCTGCCTCACGTGCGCAAGTGCGTGGTGGTGGTCTT GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCCTGGATGCCCATGCCA TCTGGCACATCAGCACCATCCCTGTCCACGTCCTCTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG GCCCTGCTGGCCTCCCCTCAACCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACTTGGACATGA AGGATGTGGGCCCAGAATCATGTGGCCAGCCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG AAGGCCTCCCAGCATCTGGGACTCGAGAGTGGGCAGCCCCTCTACCTCCTGGAGCTGAACTGGGGTGGAACTGA GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCCTCCCCACCAGCCTCCTCCCCACATCCCCAGCTG $\tt CCTGGCTGGGTCCTGAAGCCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC$ GGTTCACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTCACCCTGACC GTTGCCCTAGCCAGGTTCCCAGGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCCAGGGCA AGGATCCTGTGCTGCTGGTTGAGAGCCTGCCACCGTGTGTCGGGAGTGTGGGCCAGGCTGAGTGCATAGG TGACAGGCCCTGAGCATGGGCCTGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGT TGCGCGTGCTGGTGGGCATGTGAGATGACTGCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGAT GAGGGAATCCTGTCACCATCAATAATCACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGC ${\tt CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGCCCCTTTGCCCGCCTCCTGCAAAC}$ ACACTATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCCTGCCCCACAGG AATAGAATGGAGGGAGCTCCAGAAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTTG $\tt CTCTGCCCCTGACCCCTTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG$ GCCTGCGCTAGCTTCTTTTGATACTGAAAACTTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLLQGLSLLELLDFPPLFWVLDAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGA CTATGGCTCCCCAGAGCCTGCCTTCATCTAGGATGGCTCCTCTGGGCATGCTGCTTGGGCTG CTGATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGAC TGGATGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGG CAGGCTGTCCCTGCAGGATCCCACGTACGGCTGAATCTTCAGACTGGGGAAAGAGAGGCAAA ACTCCAATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCA ACACCTACACATCTCAGGATCTCAAGAGTGCACTGGCAAAATTCAAGGAGGGGGCAGAGATG GAGAGTTCAAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGA GGAACTGAAGAAGACTTTGATGAGCTGAATGTTGTCATTGAGACTGACATGCAGATCATGG TTTGATCTTGAATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCTTTGGTGG TCTTCAAGTGGTGATCAATGGGCTGAACAGCACAGAGCCCCTCGTGAAGGAGTATGCTGCGT TTGTGCTGGGCGCTGCCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGA GCCCTGCAGAAGCTGCTGGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGAT CCTGTTTGCACTGTGCTCCCTGCTGCGCCACTTCCCCTATGCCCAGCGGCAGTTCCTGAAGC TCGGGGGGCTGCAGGTCCTGAGGACCCTGGTGCAGGAGAAGGGCACGGAGGTGCTCGCCGTG CGCGTGGTCACACTGCTCTACGACCTGGTCACGGAGAAGATGTTCGCCGAGGAGGAGGCTGA GCTGACCCAGGAGATGTCCCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAG GCCTGTGGGAACAGGGCTGGTGCGAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGAT GCCCGTGAGAAGGTGCTGCAGACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCG TCAGGACCCCCAGCTCGGCAGGACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCA GCCTGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAAC AGCTTGCTGAAGGAGCTGAGATGAGGCCCCACACCAGGACTGGACTGGGATGCCGCTAGTGA GGCTGAGGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCT

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL
DAEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRTLVQEKGTEVLAVRVVTLLYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

Important features:

Signal peptide:

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCCTTCCC CGCGGACTCCCAGCTGGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCC $\tt CCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT$ CGCAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG TCCATGTGACCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCT GTCTCTGTCCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT ATGTTTCTGGTCTCCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCCTATTACTTCCTGACTTC AGCCTTTCTGACAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG CCTGTGAGAGGAGGCGTACTGGGCCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG GGACTGACATTCCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC TCTTGTGTAAGGAC**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG CCCATGACTGAGCCCAGCCCCAGCCCGGGTCCATTGCCCACATTCTCTGTCTCCTTCTCGTC GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCCTTTTGTGACCGTTAGTCTCTAAGCTT TACCAGGAGCAGCCTGGGTTCAGCCAGTCAGTGACTGGTGGGTTTGAATCTGCACTTATCCC CACCACCTGGGGACCCCCTTGTTGTGTCCAGGACTCCCCCTGTGTCAGTGCTCTCCAC CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC CTGGGGTGCCCTCGTGATGTCCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC GAGGTGGGGGGGGGGGGGGTATATTGGAACTCTTCTAACCTCCTTGGGCTATATTTTCTC TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT GGGGGAAAGGAAGTGCATGTTTGGGAACTGGCATTACTGGAACTAATGGTTTTAACCT CCTTAACCACCAGCATCCCTCTCCCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG ATCATTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATCAAGGTGATTGTGATTTTGACT AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIILVAGAFFWLVSLLLASVVWFILVHVTDR SDARLQYGLLIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS FGIISGVFSVINILADALGPGVVGIHGDSPYYFLTSAFLTAAIILLHTFWGVVFFDACERRR YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

FIGURE 221

FIGURE 222

GACCGACCGTTCAGATGCCCGGTTCCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTTGGGGAGTTGTTTTTTTGATGCCTGTGA
GAGGAG

FIGURE 223

NGTTGGAGAAGTGGCGGGACNTTCATTTGGGGTTTCCCCCCTTTCCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCTGCTCCTGGCCTCTGTGGTTCATCTTTGGTCCATGTGAC
CGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCCC
CAGNTGGCGCGCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGGTGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTATCATCTTGGTCCATGTGA
CCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGAGCAG TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCCACCCC TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGGTCACAGAGCATGTT CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA GGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGGTCGGATGACAGCAGCAGCCGCATCATCA ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG GAAGAAAGTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC AGCAGATGTTCCAGGGGGTCAAATCCATCCCCCACCCTGGCTACTCCCACCCTGGCCACTCT AACGACCTCATGCTCATCAAACTGAACAGAAGAATTCGTCCCACTAAAGATGTCAGACCCAT CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC TGCAGGGACTCGTGTCCTGGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTAC ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCAT CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG ACCCTCATTCCTTCCCAGAGATGTTGAGAATGTTCATCTCTCCAGCCCCTGACCCCATGTCT CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG GAACAATTTCCAAAACTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA CTGAGAAGTGGAAAAAAAAA

FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCPSAGTKCL VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGRDSCQGDSGGP VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCGGCGCCCCAACTTGAGGACCGGCCGCGA CAAGCCGCAGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC TGCTGGCTGTAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGCTGGCCTGGCTGTGC TGCTGGCTGTAGCTGTCACCGGTGCCGTGCTCTTCCTGAACCACGCCCACGCGCCGGGCACG GCGCCCCACCTGTCGTCAGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGA AAGGGCGACAGCTCGCACCTCAGCATCCTCATTGACCCGCGCTGCCCCGACCTCACCGACA GCTTCGCACGCCTGGAGAGCGCCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGC CCGGCTGCTGGCCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCATG CCAGGTGTACTGTGACATGCGCACGGACGGCGGCGGCTGGACGGTGTTTCAGCGCCGGGAGG ACGCTCCGTGAACTTCTTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCACC GGGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACACGGCTGCCTACGAGCT GCACGTGGACCTGGAGGACTTTGAGAATGGCACGGCCTATGCCCGCTACGGGAGCTTCGGCG TGGGCTTGTTCTCCGTGGACCCTGAGGAAGACGGGTACCCGCTCACCGTGGCTGACTATTCC GGCACTGCAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTCACCACCAAGGACCGTGA CAGCGACCATTCAGAGAACAACTGTGCCGCCTTCTACCGCGGTGCCTGGTGGTACCGCAACT GGTGCCTAGGGGGTGTGGGGTTCCGTTCTCCCTTCCCCTCCCACTGAAGTTTGTGCTTAAAA AACAATAAATTTGACTTGGCACCACTGGGGGTTGGTGGAGAGGCCGTGTGACCTGGCTCTCTGCCCAGTGCCACCAGGTCATCCACATGCGCAG

FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDTLADQLPRLLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRDSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASYADG
VEWSSWTGWOYSLKFSEMKIRPVREDR

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCCAGGGCAATCCGACCACATTTCACTCT CACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC TTTGTGCTTGGTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACC CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC TGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGAACCCCTTTCACTTCTGAACTG TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG TGGTGAAGCCAGAGAGCCTCCATGTCCCCCCTGAAACATTAGGCGAAGGTGAC**TGA**TTCGCC CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG TTTCCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG GCTTAGAGATAACTTTTAGCTCTCTTTCTCAATGTCTAATATCACCTCCCTGTTTTCAT GTCTTCCTTACACTTGGTGGAATAAGAAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC AGCAAATACACAAGGAATTCTTTTTGTTTGTTTCAGTTCATACTAGTCCCTTCCCAATCCAT CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG GTACTGAAGATTTAATAATAATAATGTAAATACTGTGAAAAA

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD CKELKRCVCERRAGMVKPESLHVPPETLGEGD

FIGURE 231

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGCTC CCCGAGCCCTCCGGATCCGCCCCTCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCCATGCGGGCATCGCTGCTG TCGGTGCTGCGGCCCGCAGGGCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCTGAGCCT GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGCCCAGGCCCCCCAACCTGGAGACTCTG AGCTGCCGCCGCGCGAACACCAACGCGGCGCGCCCAACTCGGTGCAGCCCGGAGCG GAGCGCGAGAAGCCCGGGGCCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTACATCAGCA CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC ACGCTGGGCGTGGACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC GGGCGCACGGGCCCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC CCATTGGACACCTGCACCTGCGCGCCACCTGCTGGAGCACGACGACGACGACTTTGAC TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG GAGAGCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGCTGCTGTCGCGCATGCTG $\tt CTGCAACAACTGCGCCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCCCCTGA$ CGAGTGGCTGGGTCGCATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT TTCCGAAGTGCCCTGACAGCCCACCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA AGCTTTCGCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT CCAGCACCATCCCGCCCGGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCACTGCGTGGGGCTGACCGGGCTG ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG CGGCTCCAGAAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCGGGGTATGGA ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCAGGGAGGCCGCCGGCCCCTCACTC GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCTGTGCCCTATGTCACT GAGGCCTCACGTCTCACTGTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCTGG CTTCTTGGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC TGCTGCTACTGTATGAGCCGCGCCAGGCCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG TGTGCAGACAGCCGCACCCTCACCACTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC TGGACACACTGTTCCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCCAGAGCTGGGCCGTGACACTGGCCGCT TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGGGCTGCTGGAGAGCCTGGATGTGTACGAGCT ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG CAACAGCACC**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCACCCCACTT CTCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC AAGCTGGCCCACTGGTCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG GGACGTGCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT GCTGATTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC CTCTGGGCCCTGGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGAGCTGAGGAGGGG GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA AGTGTGGAAAAA

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPPRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLLE
QHGDDFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCILDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALTAHPVRDPVHMYQLHKAFARAELERTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGPGPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLESLDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQEQGNST

FIGURE 234

GCTCTGGCCGGCCCCGGCTTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT AGCTATTAGCCAATTCGGCAGGGCCCGCTTTTTAGAAGCTTGATTTCCTTTGAAGATGAAAG ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG GCGGAGGCGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT TTCCCCGCCCTGAGACCCTGCAGCACCATCTGTCATGGCGGCTGGGCTGTTTGGTTTGAGC ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACTTGTATGAGAAGAACCCA GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC CTTCCCATCATGGAATCCAACTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG** CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER LVKYREANGLPIMESNCFDPSKIQLPEDE

FIGURE 236

FIGURE 237

TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTCATCACCCCGCTGCCTTCCGGGG ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTCGGAGCTTCAGCGGGAAGGAGTG TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA GCTGCACCTGTCATTCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCCACCCTTCCTGC AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA TCGCTTGTTCCACACCAGCTACCACTCCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC ATCACGGGGCAGGGAAAGAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCTCACGGA GCCCTGCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACCTACAACCAGGACA ACGAGACATTAGAGGTGCACCCACCCCGACCACTACATATCAGGACGTCATCCTAGGCACT CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACTCTCGAAACCT CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCAGTGCCCTTCCTGC ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGGAGCTGAGCACACTGCTGTAC AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAAACCAAGTTACATCC ACTACCAGCCTGCCCAGGACCGGCTGCAACCCCACCTCCTGGAGATGCTGATTCAGCTGCCG GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCCATCTGTCCTCAGCGCCCTTGTGCCCA GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCCACTC<u>TGA</u>TT CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT TCTGCCACTTGCTCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTTGAATTTGAATTAA CTTAGAAATTCATTTCCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA AGTGGTCGGTGGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACCACAGAAAGGTC GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTGTAGTGGA

FIGURE 238

MPLALLVLLLLGPGGWCLAEPPRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSISWELRQTLSVVFDAFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTVPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

FIGURE 239

FIGURE 240

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

FIGURE 241

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC TCTAGAACCCGACCCACCACCATGAGGTCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCCACACCA ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG GTGTCAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAAGTCA GCACAGAATGCTGGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCCACCCCTGCCCCTTTCCAG AGCCCCACGACGCAGAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCCAACCTCACTCTC TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT GCCGTGGTGGCCAACGGGGGCATCCTGAACAACTCCCACATGGGCCAGGAGATAGACAGTCA CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCCTGGAAGGCAC CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT TCTGGTTCAGGCACAGACCCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCATGAGCGCTTTTCTGAT CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA GAGAGAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC $\tt CCGGAACTGCCAAAGCCAAGAAC{\color{red}{\textbf{TGA}}} CCGGGGCCAGGGCTGCCATGGTCTCCTTGCCTGCTC$ CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCCTGAGGCCAGG GATTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAACACTCTT CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCCTTG TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACTGATAATAATACAAATGATTGTT

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
NQSEWDRLEHFAPPFGFMELNYSLVQKVVTRFPPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 243

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT RDCTIPAYYKRCARLLTRLAVSPVCMEDK

FIGURE 245

GGGCTGGGCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG CCCGACCCGGCCGCCCAGCCCCACCATGCCACCGCGGGGCTCCGCCGGGCCGCCG CTCACCGCAATCGCTCTGTTGGTGCTGGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCCTGTTGCTACCTGT ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTTGAAGGCCAGGAGATTCCAATGACAGGCATC CCAGTGCAGCCAGTATACCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC CAGTCTACAACCCTGCAGCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC TGAGGAACCAGCCATGTCTCTGCTGCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA ${f ACTATGAGGGGTTGGGGGGGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG$ GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTFCCGTCYHRYC CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQLQSP FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

FIGURE 247

GGGGGAGCTAGGCCGGCGCAGTGGTGGTGGCGCGCGCGCAAGGGTGAGGGCGGCCCCAGAA TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG AAGAACCCCTGCAGGTCCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAG CAACACAATTTGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTC AAAGCAAGTTTCTCAATCAAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCC ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA AGATGAGCACCTATCTGGTGGCCTTCATCATTTCAGATTTTGAGTCTGTCAGCAAGATAACC AAGAGTGGAGTCAAGGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTCAGCATACCGTATCCCC TACCCAAACAAGATCTTGCTGCTATTCCCGACTTTCAGTCTGGTGCTATGGAAAACTGGGGA CTGACAACATATAGAGAATCTGCTCTGTTGTTTGATGCAGAAAAGTCTTCTGCATCAAGTAA GCTTGGCATCACAGTGACTGTGGCCCATGAACTGGCCCACCAGTGGTTTGGGAACCTGGTCA CTATGGAATGGTGGAATGATCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTG TCTGTCAGTGTGACCCATCTTGACTGAAAGTTGGAGATTTTCTTTGGCAAATGTTGGAGTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTTGACGCAAATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATG CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAAGTGGTATTGTACAGTATCTCCAGAAGCA TAGCTATAAAAATACAAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG ATGGTGTAAAAGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT TCCCCTAATAACCATCACAGTGAGGGGGGGGAATGTACACATGAAGCAAGAGCACTACATGA AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTCATCACC AGCAAATCCAACATGGTCCATCGATTTTTGCTAAAAACAAAACAGATGTGCTCATCCTCCCAGAAGAGGTGGAATGGATCAAATTTAATGTGGGCATGAATGGCTATTACATTGTGCATTACG AATGATCGGGCAAGTCTCATTAACAATGCATTTCAGCTCGTCAGCATTGGGAAGCTGTCCAT TGAAAAGGCCTTGGATTTATCCCTGTACTTGAAACATGAAACTGAAATTATGCCCGTGTTTC GACAGACGAGGCTCAGTCTCAGAGCAAATGCTGCGGAGTGAACTACTACTCCTCGCCTGTG AATGGAAACTTGAGCCTGCCTGTCGACGTGACCTTGGCAGTGTTTGCTGTGGGGGCCCAGAG CACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTTTTCTTTGTCCAGTACTGAGAAAA GCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGAAAAGCTTCAATGGCTACTAGAT TACAAAAGTTTGAACTTGGCTCATCTTCCATAGCCCACATGGTAATGGGTACAACAAATCAA TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTGAAGAAAACATCGGTTGGATGG ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAA**AAA TTCCTCCCTTGCCCGGTTCCTGTTATCTCTAATCACCAACATTTTGTTGAGTGTATTTTCAA ACTAGAGATGGCTGTTTTGGCTCCAACTGGAGATACTTTTTTCCCTTCAACTCATTTTTTGA CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTCATGAATGGGCTA

FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE HPPQEQIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA ARMAFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA FIISDFESVSKITKSGVKVSVYAVPDKINOADYALDAAVTLLEFYEDYFSIPYPLPKODLAA IPDFOSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL WLNEGFAKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD DVSYDKGACILNMLREYLSADAFKSGIVOYLOKHSYKNTKNEDLWDSMASICPTDGVKGMDG FCSRSQHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL TGLLKGTHTAVSSNDRASLINNAFOLVSIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP MYKLMEKRDMNEVETOFKAFLIRLLRDLIDKOTWTDEGSVSEOMLRSELLLLACVHNYQPCV ORAEGYFRKWKESNGNLSLPVDVTLAVFAVGAOSTEGWDFLYSKYQFSLSSTEKSQIEFALC RTQNKEKLQWLLDESFKGDKIKTQEFPQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS SSIAHMVMGTTNOFSTRTRLEEVKGFFSSLKENGSOLRCVOOTIETIEENIGWMDKNFDKIR VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

 $\texttt{CAGCCACAGACGGGTC} \underline{\textbf{ATG}} \texttt{AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC}$ TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCCGGCCTCTCCCTGATC TTGGGCCCCACAGCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGG AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGA ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT GTGTCAGGAGACGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCTCCTGGG GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG CAGCGTTCTGCTGAACTCCCTCCCTCAAGCTGCCCCTGTCCCAGGAGACCGGCAGTGTC CTACCTGTGTGCAGCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG GGCGCCACTCATTGTTATGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAAT GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG GTGGGGAGTGGTTTGCCCTTCCTGC**TAA**CTCTATTACCCCCACGATTCTTCACCGCTGCTGA CCACCCACACTCAACCTCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC
PSC

FIGURE 251

CAGGATGAGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCGGACGGCCTGGAAGAGTCG GCCCACGGGAGAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCGTCAT GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTCGCC GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA CGCCCAGCTGTCCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCGTGTCTTCATCGGCATCAAC GACCTGGAGAAGGAGGCCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA GTGGCGCAGCGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTGGGGGCCCCACATGTCCCTGCAGGGTT GGCAGGGACAGACCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG AAAGTGTTCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC

FIGURE 252

MRGNLALVGVLISLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG PTGEKGDMGDKGQKGSVGRHGKIGPIGSKGEKGDSGDIGPPGPNGEPGLPCECSQLRKAIGE MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS GGWNDVACHTTMYFMCEFDKENM

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG CACCAGTGTGTGAGGGGAGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC $\texttt{CAGCCCCGGCACAGAGCTGCTCCACAGGCACC} \underline{\textbf{ATG}} \texttt{AGGATCATGCTGCTATTCACAGCCAT}$ TTCCTGGCGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCC CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC TTTA<u>TAA</u>GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT CCTACATTAAAAATATAATGTCTCTCTCTATTCCTCAACAATAAAGGATTTTTGCATATGAA

FIGURE 254

MRIMLLFTAILAFSLAQSFGAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK ALSQASTDPKESTSPEKRDMHDFFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCGCTGTGCCTGTGCC CGCGCTGTCGCCGCTGCTACCGCGTCTGCTGGACGCGGGGAGACGCCAGCGAGCTGGTGATTG GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC $\verb|CCATAGTTGCTGCAGGAGTGGAGCC| \textbf{ATC}| AGCTGCGTCCTGGGTGGTGTCATCCCCTTGGGGC| \\$ TGCTGTTCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTTTAGAGGAG CTGCTCAGCAAATACCAGCACAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG AAACCT**TAG**ACTCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACTACAG GCATGCACCATGGTGCCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA ACAACACGTGGGTTCCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC TCCACTGGGAACACAGCTCTCAGCCTTTCCCACCTGGAGGCAGAGTGGGGAGGGGCCCAGGG CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG TGGAAACTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAAATC CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG GTGGAGGTGGGGGGGGGGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLEELLSKYQHNESHSRVRRAIPREDKEEILML HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

FIGURE 257

FIGURE 258

 ${\tt MGSGLPLVLLITLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTS} \\ {\tt VTLHHARSQHHVVCNT}$

FIGURE 259

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFSVENECLVDLCLLRICYKLSGVPNQCRVPLP SDCSK

Important features:

Signal peptide:

amino acids 1-29

FIGURE 261

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGGAGCACCACCGGAGCCCTTGAGACATCCTT GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTGCAGG**ATG**ATGGTGGCCCTT CCAGGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGG AAAAATGTACCCAAGCAACGAGGGCATACATTCAAGAATTCCAAGAGTTCTCAAAAAATATA TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTT ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC TTTGAAAATAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC ATACGGGCATTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC CTGGCAGGGAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTCATAACCAAGCAACTT CTAATGAGATAATCAAATATAACCTGCAGAAGAGGACTGTGGAAGATCGAATGCTGCTCCCA GGAGGGGTAGGCCGAGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT GGATGAGCATGGGCCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAG GATGCTGAAGCCTCATTCCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC CCAACTTGTTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCCAGAGAT AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAAACTCCAGACAAAGAGAAAA ${\tt GCTGCCTCTGAAG}$ ${\tt TAA}$ ${\tt TGCATTACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTC}$ AGTGTGTAGAAGTGGAAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAA CCTCCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTTTTACT CCCTAATATTCACCACTGGCTTTTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCAACTAAAATACTATTAATATTTCTTT CTTTTCTTTTTTTTTTGAGACAAGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACTCC AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC CACACCTGGCTTAAAATACTATTTCTTATTGAGGTTTAACCTCTATTTCCCCTAGCCCTGTC CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT CCAGGTGTGGAGTGTTTGCACATCATTGAATTCTCGTTTCACCTTTGTGAAACATGCACAAG TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT TGTTCACCTACTCTTATAGTCAATGCGTTCATCGTTTCAGCCTAAAAATAATAGTCTGTCCC TTTAGCCAGTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCCTCTGT

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFRVLEQGLEKCTQATRAYIQEFQE FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL QEAEEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNSPKVYLLIGSRNNTV WEFANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDT PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 263

GGGCGCCCGCGTACTCACTAGCTGAGGTGGCAGTGGTTCCACCAAC**ATG**GAGCTCTCGCAGA TGTCGGAGCTCATGGGGCTGTCGGTGTTGCTTGGGCTGCTGGCCCTGATGGCGACGGCGGCG AAATGGATTTCCACCTGACAAATCTTCGGGATCCAAGAAGCAGAAACAATATCAGCGGATTC GGAAGGAGAACCTCAACAACACACTTCACCCACCGCCTCCTGGCTGCAGCTCTGAAGAGC CACAGCGGGAACATATCTTGCATGGACTTTAGCAGCAATGGCAAATACCTGGCTACCTGTGC AGATGATCGCACCATCCGCATCTGGAGCACCAAGGACTTCCTGCAGCGAGAGCACCGCAGCA TGAGAGCCAACGTGGAGCTGGACCACGCCACCCTGGTGCGCTTCAGCCCTGACTGCAGAGCC TTCATCGTCTGGCTGGCCAACGGGGACACCCTCCGTGTCTTCAAGATGACCAAGCGGGAGGA TGGGGGCTACACCTTCACAGCCACCCCAGAGGACTTCCCTAAAAAGCACAAGGCGCCTGTCA TCGACATTGGCATTGCTAACACAGGGAAGTTTATCATGACTGCCTCCAGTGACACCACTGTC ACACGCTGCTGTATCTCCCTGTGGCAGATTTGTAGCCTCGTGTGGCTTCACCCCAGATGTGA AGGTTTGGGAAGTCTGCTTTGGAAAGAAGGGGGGAGTTCCAGGAGGTGGTGCGAGCCTTCGAA CTAAAGGGCCACTCCGCGGCTGTGCACTCGTTTGCTTTCTCCAACGACTCACGGAGGATGGC TTCTGTCTCCAAGGATGGTACATGGAAACTGTGGGACACAGATGTGGAATACAAGAAGAAGC AGGACCCCTACTTGCTGAAGACAGGCCGCTTTGAAGAGGCGGCGGGTGCCGCCGTGCCGC CTGGCCCTCTCCCCCAACGCCCAGGTCTTGGCCTTGGCCAGTGGCAGTAGTATTCATCTCTA CAATACCCGGCGGGGCGAGAAGGAGGAGTGCTTTGAGCGGGTCCATGGCGAGTGTATCGCCA ACTTGTCCTTTGACATCACTGGCCGCTTTCTGGCCTCCTGTGGGGACCGGGCGGTGCGGCTG TTTCACAACACTCCTGGCCACCGAGCCATGGTGGAGGAGATGCAGGGCCACCTGAAGCGGGC CTCCAACGAGAGCACCCGCCAGAGGCTGCAGCAGCTGACCCAGGCCCAAGAGACCCTGA AGAGCCTGGGTGCCCTGAAGAAG**TGA**CTCTGGGAGGGCCCGGCGCAGAGGATTGAGGAGGAG GGATCTGGCCTCCTCATGGCACTGCTGCCATCTTTCCTCCCAGGTGGAAGCCTTTCAGAAGG AGTCTCCTGGTTTTCTTACTGGTGGCCCTGCTTCTTCCCATTGAAACTACTCTTGTCTACTT AGGTCTCTCTTCTTGCTGGCTGTGACTCCTCCCTGACTAGTGGCCAAGGTGCTTTTCTTC CTCCCAGGCCCAGTGGGTGGAATCTGTCCCCACCTGGCACTGAGGAGAATGGTAGAGAGGAG AGGAGAGAGAGAGAATGTGATTTTTGGCCTTGTGGCAGCACATCCTCACACCCAAAGAAG TTTGTAAATGTTCCAGAACAACCTAGAGAACACCTGAGTACTAAGCAGCAGTTTTGCAAGGA TGGGAGACTGGGATAGCTTCCCATCACAGAACTGTGTTCCCATCAAAAAGACACTAAGGGATT TCCTTCTGGGCCTCAGTTCTATTTGTAAGATGGAGAATAATCCTCTCTGTGAACTCCTTGCA AAGATGATATGAGGCTAAGAGAATATCAAGTCCCCAGGTCTGGAAGAAAAGTAGAAAAGAGT AGTACTATTGTCCAATGTCATGAAAGTGGTAAAAGTGGGAACCAGTGTGCTTTGAAACCAAA TTAGAAACACATTCCTTGGGAAGGCAAAGTTTTCTGGGACTTGATCATACATTTTATATGGT TGGGACTTCTCTCTCGGGAGATGATATCTTGTTTAAGGAGACCTCTTTTCAGTTCATCAAG

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASCG
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG CAGTGTTTTGCCTTCACCCCAAGTGACC**ATG**AGAGGTGCCACGCGAGTCTCAATCATGCTCC TCCTAGTAACTGTGTCTGACTGTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT GGGGCGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA AGCACCACACCTGTCCTTGCTCCAACCTGCTGTGCTCCAGGTTCCCGGACGCAGGTAC CGCTGCTCCATGGACTTGAAGAACATCAATTTT**TAG**GCGCTTGCCTGGTCTCAGGATACCCA CCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC TCTCCCAGTCCCTACACTGACTACCCTGATCTCTTTGTCTAGTACGCACATATGCACACAG GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTCCCTGCTCAGGCTGCCAGAGAGGTGGTA AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGT GCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA GGGAGGCCAATCAGCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACTCCTTTCCA AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTCAGACCAGGGAGG CAACTACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAA

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA**ATG**TCTTTC CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC TGCTGCCTTGGCATTGCTGCCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA GCAAAGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT CCTCTGTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAGTTCTATGGTCCACAAAGGCGGAAGAT GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACTTCTTCCGAGCCTGGAACGGAG GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGATCA GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAA**TGA**T TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTCATGGGATGTATT GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA CAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC TCTACTAAAAATACAAAAATCACCCGGGTGTGGTGGCAGCACCTGTAGTCCCAGCTACCCG GGAGGCTGAGGCAGGAGATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA

FIGURE 268

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKAALEYLEDIDLKTLEKEPR TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEGFILGGVFV VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 269

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSTSSSSDGNLRPEAITAIIV VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

FIGURE 271

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK GIVKGRNLDSRGLILGAEAWGRGVKKNT

FIGURE 273

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGTTATTCAGAGGTTTTGTTTTCCTCTTAGTTCT GTGCCTGCTGCACCAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA TGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCT TCTACTTGGAAAAAACAAAATGAATATGGACCACCAGGCAAACTGTTTGTCCATGAGTGGG CTCACCTCCGGTGGGGAGTGTTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACTGTATG GAAAAGATTGTCAATTCTTTCCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG CAAAGTATTGATTCTGTTGTTGAATTTTGTAACGAAAAAACCCATAATCAAGAAGCTCCAAG CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATT
TTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC
AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT
AAATCGAATGAATCAAGCAGCAAAACATTTCCTGCTGCAGACTGTTGAAAATGGATCCTGGG
TGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAAATAAAAAGC AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCAT ${ t CTGCTCTGGAATTAAATATGCATTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT}$ CCGAAGTACTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG AAACAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCCAGAAGTCCCTT CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA TTTCTCTCTGGGATCCCAGTGGAACAATAATGGAAAATTTCACAGTGGATGCAACTTCCAAA ATGGCCTATCTCAGTATTCCAGGAACTGCAAAGGTGGGCACTTGGGCATACAATCTTCAAGC CAAAGCGAACCCAGAAACATTAACTATTACAGTAACTTCTCGAGCAGCAAATTCTTCTGTGC CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT TGAATCACAGAATGGACATACAGAAGTTTTGGAACTTTTGGATAATGGTGCAGGCGCTGATT CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT AGCTTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCACT GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTCAGCCGAACAGCATCCGGA GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCCAACCC ACATATTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT GCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC TACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAGTACCACCATT**TGA**ACCTTA ACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTAAAAAAACAAAACAATGTAAGT AAAGGATATTTCTGAATCTTAAAATTCATCCCATGTGTGATCATAAACTCATAAAAATAATT TTAAGATGTCGGAAAAGGATACTTTGATTAAATAAAAACACTCATGGATATGTAAAAACTGT CAAGATTAAAATTTAATAGTTTCATTTATTTGTTATTTGTAAGAAATAGTGATGAAC AAAGATCCTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT

FIGURE 274

MGLFRGFVFLLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTY
LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAAKHFLLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTDGEDNTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDGVYSRYFTAYTENGRYSLKVRAHG
GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRTASGGAFVVSQV
PSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP
OANPDDIDPTPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 275

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG GGCAGGGGTGACAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAG ACGCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACC CCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTT GTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTCCGGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGG GCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGCTTAGCCCAGGGAAGCCCTTCCGTGGGG GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGCTGATTCCTGAATGATGATCGCC GGGGGCTGCTTGCGTGGATTTCCCGGGTGGTGTTTTGCTGGTGCTCCTCTGCTGTGCTATCTCTGT CCTGTACATGTTGGCCTGCACCCCAAAAGGTGACGAGGAGCAGCACCGCAACTACGTGA CCCACGGGGAAGGAGGGGTACCAGGCCGTCCTTCAGGAGTGGGAGGAGCACCACCGCAACTACGTGA GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGCTCAGGAA TGGGCAGTACCAGCGATGCTGGCTGGGTCTGGACAGGAGCCCCCCAGAGAAAACCCAG GCCGACCTCCTGGCCTTCCTGCACTCGCAGGTGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG CCACAGAGTATGCAGCAGTGCCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGG CTTCGCAAACAGAAACAGAAGACAGTAGCAAAAAAACA<u>TGA</u>ACTCCCAGAGAAGGATTGTGGGAGA CACTTTTTCTTTCCTTTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGACG ACAAAAGAATTGGACTGATGGGTCAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTTAC GTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTCAGGCACTGAAT GTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA GAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAA

FIGURE 276

MMMVRRGLLAWISRVVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIES
ALETLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPI
MKVKNEKLNMANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR
LNTQPGKKVFYPVLFSQYNPGIIYGHHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

FIGURE 277

GAAAGA**TG**TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC CTGGAATTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT CAGGGATCTGGCAACGTAGAAGAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA CCAAGAGCAGATCATATATTTTGTTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCT TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC TCAAAATATTCTAAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG TAGTTATTGATTTAAGCATTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATT TTATCCTGTTATCACACCAACAGTTGATTATATATTTTCTGAATATCAGCCCCTAATAGGAC AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG

FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVÄFSMRKVP NREATEISHVLLCNVTQRVSFWFVVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLE FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC ENMITIENGIPSDPLDMKGGILMMPS

FIGURE 279

AACTCAAACTCCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATG**T ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATACAGCTCACAGCTCTTTGG CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC CAACCCATGAGTGGCCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGAAAAGGCTCAACCAAGAGAAAAAAGGTCTCT GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA GAACCCTAGTATTTCTTGAAGTTAATGGAAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA GTGTAAATTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTGCCCTTAA GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC AATTTGTCTGTTACATTTCCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA ATGTGTTTACTCTTTCCTTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG TTTCTGATTAACAGTAAATCCTAAATTCAAACTGTTAAATGACATTTTTATTTTTATGTCTC TTTGTCG

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT WNFRPLDGGPEQFVFYYHIDPFQPMSGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFQHYRKKRWAER AHKVVEIKSKEEERLNQEKKVSVYLEDTD

FIGURE 281

FIGURE 282

 ${\tt MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTTAT}$ ${\tt TAASTTARKDIPVLPKWVGDLPNGRVCP}$

FIGURE 283

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

FIGURE 285

FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ IQERLHTAALPA

FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACTGTGTCCTGTTCTCCCAATCAGCTCAG GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC TAAA**TAG**GTCTTTCTCCAATGTGTCCTCCAAGCAAGATTCATCATAACTTATAGGTTCATGA TCTCTAAGATCAAGTAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCCTACATCAGAGACTCTAGGT GCTATATAATCCAAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT TACCCTTTTTTTGGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT CAATTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCCTTTTTTACATTT TCGTATATTTATTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT TGGAAGCTGAAAACTGAATTTAAAGAATGCTATCTTGGAAAATTGCATACGTCTGTGCAATT TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTAGAAGCAAACAATTTT AAATATATTTTGTTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT CCCACTTTGCAAACTTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTCATTGCTCA

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGDLKTQIEKLWT EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQGK WSDEACRSSKRYICEFTIPK

FIGURE 289

FIGURE 290

MKLAALLGLCVALSCSSAAAFLVGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPLKLLLS SLGIPVNHLIEGSQKCVAELGPQAVGAVKALKALLGALTVFG

FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT CCTTGGCCTCCGCAGCCGATCAC ATG AAGGTGGTGCCAAGTCTCCTGCTCCTCCTGGCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCCAGCCCC TCAGAACCAGACCAGCAGGTAGTGCAGGCTCCCAGGGAGGAAGAGAAGATGAGCAGGAGG CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGCAGCAGCTT GCCAAGGAGACTTCAAACTTCGGATTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG CAACATGGTCTTCTCCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCA CAGGGCCGACTGAAACCCAGATCAAGAGAGGGCCTCCACTTGCAGGCCCTGAAGCCCACCAAG CCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACT GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCGTGCCTATGAATTTTCGCAATGCCTCA CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTCCCAAACT GTTTGATGAGATTAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGA AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA TTTTCGTTGTCATGTCCTCAAACTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCCTCA TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCTTTG CTGACCTTAGTGAACTCTCAGCTACTGGAAGAATCTCCAAGTATCCAGGGTTTTACGAAGA ACAGTGATTGAAGTTGATGAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTAC TGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTCATTTCATGATCTATGAAG AAACCTCTGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTA**TAA**TTCAGG TACCAGCAATGGATGGCAGGGGAGAGTGTTCCTTTTGTTCTTAACTAGTTTAGGGTGTTCTC AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGGATACATTCAAAGACCCCCAGCAGATGC AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACTGATTATAGAGAAGGCTA CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAAGATTCA CATCCTGGGTGGACAGAGCAGGACGATGCAAGATTCCATCCCACTACTCAGAATGGCATGC TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTCATTTAATGTTTTTTGGACCATGGT TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAA

FIGURE 292

MKVVPSLLLSVLLAQVWLVPGLAPSPQSPETPAPQNQTSRVVQAPREEEEDEQEASEEKAGE
EEKAWLMASRQQLAKETSNFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQI
KRGLHLQALKPTKPGLLPSLFKGLRETLSRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYF
DTECVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLTPFDP
VFTEVDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHL
ALEDYLTTDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFSPFADLSELSA
TGRNLQVSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLF
LGRVVNPTLL

FIGURE 293

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGACATGAG
GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTCCCAG
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG
GCCTGGGGCCCCGTGTGGTGGAGCCTCCGGAGAAGGACCAGCTGGTGGTGCTGTTCCC
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGCAGGGCCCCA
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCC
GAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGACCAGGGCGAGGAGAGCCACA
TCTACCACCCCCAGTAGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA
TCTACCACCCCCAGTAGGCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAAGGCCCAGG

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE RPRLWVMPNHQVLLGPEEDQDHIYHPQ

FIGURE 295

TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGC TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCT GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG CGTGGGAAGTGCACGGTGGCCGATCGCTGGTCCAGTCAGCAGGCCAAAGCAGACTACCC AGAGGGGGACGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTTGGC GACGCCCAGAAAACAGCATCTTATTACTCACCCTATGGCCAGCGGGAATTCACTGCGGGATT TGTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT CCCCAGCAGTGTGGAGATTTTTCTGGTTTTGATTGGAGTGGATATGGAACTCATGTTGGTTA CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTG GGAGGGAACCCAGACCTCTCCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCA GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI
YQKYPVKYGEGKCWTDNGPVIPVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL
FYR

FIGURE 297

GCGGAGCCGGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC CACGAGGCTGCCGCATCCTGCCCTCGGAACA**TG**GGACTCGGCGCGCGAGGTGCTTGGGCCG CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGGCCCCATGAAAGCGCAGCC ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACTCCAGTGCTAACTCAAC AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACTATGCAT TCTGAAGCAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCT CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT GGGGTGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA TGCCATCTGGGCATACAAATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGT AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA CAGAAATTATACAATCAAACTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPSDH TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQN TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIG CKMYYSRRGIRYRTIDEHDAII

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCCAGCGATGGCGACCCTGTGGGGAGGC ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA ATGCAAATATGAAGAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG CATGTTGTCCTCAGC**TAA**TTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTCACCAACT ATAATAGAGACATTTTTAAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG TGACTTTTACTAATAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA AGCACTCTCTTTTTCACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTTG TTGTTGTTGTTTGTTTGTTTGTTTGGTGGGAGGGGAGGGATGCCTGGGAAGTGGTT AACAACTTTTTCAAGTCACTTTACTAAACAAACTTTTGTAAATAGACCTTACCTTCTATTT TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG ACTTTTGCACTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT CTAAAATGCCTGGTGGCTTTTCACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG GTGTGTGTGTGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT TGCAATAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATACACATTTG TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTTGAAACATCAA

FIGURE 300

MATLWGGLLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENSGHIYNKNIS QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIYLSILGLLLLYMVYLTL VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLARSRSRANVLNKVEYAQQRWKLQVQEQ RKSVFDRHVVLS

FIGURE 301

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT PGARFQRSHLAEAFAKAKGSGGGAGGGGSGRGLMGQIIPIYGFGIFLYILYILFKVSRIILI ILHQ

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT GGATTTGAAAGTTGAGAGCAGCATGTTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG GGTGATTCAGCTCTGATGGGATGTTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT AGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC AATGATGGCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA GGAGATTGTATTTCGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG CTTCAAGGAGTGAGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCT GGTGTTCAAGAAACCATTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC ${\tt CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC}$ TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGG GAATGCCAAAAACACAGCAAGCCTTT**TGA**GAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG TGGAGACTCTCTCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC CCAGCTGTCCTCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA AAAACCAACCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHFQNRVNLVGDIFRNDGSIMLQGVRESDGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVTPAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNSLEKKSGGGMPKTQQAF

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG AAACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG GCCCTGCATCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA TGAGAATGAAAATCGCACAGGAACTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAAC TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAG TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGA TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT TAATGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT ATGAATGCATCAGTAGCTGAAAAAAAAAAAAAA

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYL QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLTWE ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFL EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

FIGURE 307

CCCACGCGTCCGCGCAGTCGCCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA GCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG CGTCGTGGCCATGCCGGCGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACTCTTCGGCTCCAAGAAGAGGCGCAG AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC **GTAG**CCAGTGAGGGCAAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFGSKKRRRRP EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA MNSEGYLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGGAGCCCTGTCTTACTGAACCTGGG CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAAACTCGAGGGTGGGATCCACTGAGG ${\sf AGTACATAGGCTGCTGGATCTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT}$ CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG CGCAGCGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG CTGTCCAAGGTGCGACTGTGCGGGGGGGGGCCGGCCGGGCCGGACCGCGGCCCGGAGCCTCA GCTCAAAGGCATCGTCACCAAACTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG ACGGAAGCATCCAGGGCACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC TGAGGGACTGCTCTACAGTTCGCCGCATTTCACAGCTGAGTGTCGCTTTAAGGAGTGTGTCT TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTACCGCCAGCGTCGTTCTGGCCGGGCC TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCCTGAAATGTAGTCCCTGGACTG GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACACCTGT

FIGURE 310

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR GPEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

FIGURE 311

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRIFGLKKRRLR RQDPQLKGIVTRLYCRQGYYLQMHPDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSKTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 313

ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG CTCCCCACCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTT $\verb|TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAA|$ AGAACTGGTGTGTGTGTTTTCCTTTTTTGAATTTCCCACAAGAGGAGAAATTAATAATACATCTGC ${\tt CAGTTGGATTTGTCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT}$ TTTTTAAATTTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCT GGATGTTGCTGTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA**ATG** TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCT GCTTGTGGTGCTGCTGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT GCTCCTGCAGCAACCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC ACCAACACACGGCTGCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG GCACTTGGAAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA ACCTCAACACTCTGGAACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTT GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACT TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAG $\tt CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACT$ GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACAT TTACATCACAACCCTTGGAACTGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC GAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT GAGCTGAAATGTCGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATA CAGGCATGTACACATGTATGGTGAGTAATTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG GACCACAGATAACAATGTGGGTCCCACTCCAGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACAC CACAGAGCACAAGGTCGACAGAGAAAACCTTCACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT GGTCATTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTA TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCCATGCCTATCGAGCATGAG CAGTTCAGTGCATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATC<u>TAA</u>AACA TTTACAGAGTTACAAAAAACAACAATCAAAAAAAAAGACAGTTTATTAAAAATGACACAAATGACTGGGCTAA TGATCTAAAGCAGACAAAAA

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

FIGURE 315

GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCCCAGCTCGCCCGAGGTCCGTCGGA GGCGCCCGGCCCCCGGAGCCAAGCAGCAACTGAGCGGGAAGCGCCCGCGTCCGGGGATC GGGATGTCCCTCCTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA CACTGAGATCAAGAGGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG AGTGGCCTTTGCTTCCAATTTCCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT ACTGGCAGCGAATCCGAGAGAAAGAGGGAGGAGGATGAACGTCTGCCTCCCAAATCTAGGATT GACTACAACCACCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA CCAGTGCACAGCAGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT ACGGTCTGAATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATT ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC TATGAAAGGAGACAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA CCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAAACCAGAATTGCAAGACTGG TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 317

 ${\tt CCATG}{\tt GCGCTCCTGCTGCTGCTGCTGCTGCGGGGTAGTGGATTTCGCCAGAAGTTTGAGTATCACTACT}$ CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA $\tt CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT$ ATTCTGGAGACAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACTGTCAGATATTGGCACATATCAGTGCAAAGTGAA AAAAGCTCCTGGTGTTGCAAATAAGAAGATTCATCTGGTAGTTCTTGTTAAGCCTTCAGGTGCGAGATGTTACG TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG TATGAGTGGCAAAAATTGTCTGACTCACAGAAAATGCCCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT TTGCTTGCTCTAGCGCTCATTGGTCTTATCATCTTTTGCTGTCGTAAAAAGCGCAGAGAAAAAAATATGAAAA GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG GCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA GTACCAAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGACTCTCCCACCTGCTAAGTTCAAGTACCCTTACAA GACTGATGGAATTACAGTTGTA**TAA**ATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTT AGGCCTCTAGTAAAGACTTAAATGTTTTTTAAAAAAAAGCACAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC ATCTACTTTATGCAATGGCATTAGACATGTAAGTCAGATGTCATGTCAAAATTAGTACGAGCCAAATTCTTTGT TAAAAAACCCTATGTATAGTGACACTGATAGTTAAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA ATTTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTCAA AGGAAATTTTAAAATTCTTACGTTCTGTTTAATGTTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG TAAAGTAAATTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGC ACAATATATAAATACTCAAGTCCAATATTAAAAACTTAGGCACTTGACTAACTTTAATAAAAATTTCTCAAACTA TATCAATATCTAAAGTGCATATATTTTTTAAGAAAGATTATTCTCAATAACTTCTATAAAAATAAGTTTGATGG TTTGGCCCATCTAACTTCACTACTATTAGTAAGAACTTTTAACTTTTAATGTGTAGTAAGGTTTATTCTACCTT TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTTTTCTTCCC TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCTGGCTT TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGGTGGAAATATAAGTTCTGAAATCTGTAGGGAAGAAGAACACAT TAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT GAGATCACATCTAAGTGACACTCCTATTGTCAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCCTAAACAA CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGGTTT CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT AGCAAACTAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT TATGAACACAAAGAAGGAAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAAGGAAAAGGAGCA GAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK APGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIAGAIIGTLL ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247, 245-251, 296-302

Myelin P0 protein.

amino acids 96-125

FIGURE 319

 ${\tt CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGC}$ CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG AGAGATGAATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG CTTCTCTGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT GAAGGGGGTTGTTCCCCAGAAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACTGTTTTCAAAAA CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAACT TTGTTCTCATCGTGTCACAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC AGTGCACACAGGCGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC TCTGAATGTCTAGACCAGGACCTCCCTCCCCTGGCACTGGTTTGTTCCCTGTGTCATTTCA AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATTC TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTACTACAACTCTATTT AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTGGCCTGTCTTTGGATT GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTCATT TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA CTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMQMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLVHTLLEFYLKTVFKNH HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL TKALGEVDILLTWMQKFYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

FIGURE 321

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ RQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT AGAATAACATCCACTTTGCCTTTCTCCCACAGGTGTCCACTCCCAGGTCCAACTGCACCTC GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACGATGTTGGGGGGCCCGCCT CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA TGAGCAGAAGATACCTCTGCATGGATTTCAGAGGCAACATTTTTGGATCACACTATTTCGAC CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC CACCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACC CCCATACCACGGCGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT GCTGAAGCCCCGGGCCCGGATGACCCCGGCCCCGGCCTCCTGTTCACAGGAGCTCCCGAGCG CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC ACGCACGCTGGGGGAACGGCCCGGAAGGCTGCCGCCCTTCGCCAAGTTCATC<u>TAG</u>GGTCG CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARNSYHLQIHKNGHVD GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY DVYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSE RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFA KFI

Important features of the protein:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

(

FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG** GAGCTCTGCAGGCCCCAGCACCCGCAGAGCAGACACTGCGATGACAACGGACGACACAGAAG TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC ACCCGTCACGGGCCTCAGAGAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACTTC CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC GAGGCCGACCACAGACGTGAGTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG TGGCTTCCCCGGAAGACCTCACTGACCCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC CGGGAACTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG GGCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG AAGGGCAGCATGTCCAAGCCCCTAACCCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC CGGAGTGTATGTATGGGGAGGGCTTCACCTGTTCCCAGAGGTGTCCTTGGACTCACCTTGG CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA TTAAAATTCACTCAGTGTGGCCCAAAAAAA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLSA
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAIFDTLCTDDSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTSAKTTMKPOOPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site. amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169, 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339, 404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 327

GCGGAGCATCCGCTGCGGTCCTCGCCGAGACCCCCGCGGGATTCGCCGGTCCTTCCCGCGG GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCCTCTCGACGCCA CCTTGACCTTTGAAGACCAAAACTAAACTGAAATTTAAA**ATG**TTCTTCGGGGGGAGAAGGGAG CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC AAAAAACATATCAGGGGACAAAGCATGTAACTTGATGATCTTCGACACTCGAAAAACAGCTA GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG CCAAGAGTTACCCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCCTCTGATCAAGAAA TAGCTCATCTGCTGCCTGAAAATGTGAGTGCGCTCCCAGCTACGGTGGCAGTTGCTTCTCCA CATACCACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC AGCCTCCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA AACCATACCGTTTACAGAAATCTCCAACTTAACTTTGAACACAGGGAATGTGTATAACCCTA CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAAACTGCTTCCTGGGAAGGT AGGGAGGCCAGTCCAGGCAGTTCCTCCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCCTGTTCCTGGTGATAGGCC TCGTCCTCCTGGGTAGAATCCTTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT ${\tt TATTTGATCAATGGGATCTATGTGGACATC} {\tt TAAT}{\tt GGATGGAACTCGGTGTCTCTTAATTCATT}$ TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG GCTCTGTTGCCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA CCACACCTGGGTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG GTCTCAAACTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTTGGTTTTTTGAGAAGGAATGAAGTG GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAAATATTAGCTAAAAAACAAAGCTCT ATGTAAAGTAATAAGTATAATTGCCATATAAATTTCAAAATTCAACTGGCTTTTATGCAAA GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC TGTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTTATATGGATTCCTTTAAAACTTATT CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGIRGNEPVYTSTQED CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP TDSKGSLETIPFTEISNLTLNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321



FIGURE 329

 $\tt CTCCCACGGTGTCCAGCGCCCAGA{\color{red}{\textbf{ATG}}} CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT$ CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT GGGATCCTCTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA ACCTCACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG TCTTTACTGATCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCTCCCCTTCTCCCAC CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG GCTGAGGCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCATGCAGC TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGAGCCTTCTGTCAGC CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC CCTTCCCAGGCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA GCTGGGCTTCTCGAAGTTTGTCTCAGCG**TAG**GGCAGGGCCCTCCTGGCCAGGCCAGCAGT GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGCATGTTCCA GCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC TGGCGTCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG GACTCTGAATTCTAACAATGCCCAGTGACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTG ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTGAAAGTA GCACAACTACTATTTTTTTTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT TTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGAC CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG TCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAAACTTGGAAAG ATGGAGGAGAAAAGGAAAAGGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT TATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTCACATAATTTGCCGGTGTTCTT TTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC GCTGCATAAAAAAAAAAAAAA



FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196</pre>

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128